METHODS OF DIAGNOSING NEUROPSYCHIATRIC DISORDERS

Cross Reference to Related Applications

This is a continuation-in-part application of United States Patent Application No. 09/236,134, filed January 22, 1999, which claims the benefit of priority under 35 U.S.C. § 119(e) of United States Patent provisional Application No. 60/078,044, filed on March 16, 1998 and of United States Patent provisional Application No. 60/088,312, filed on June 5, 1998 and of United States Patent provisional Application No. 60/106,056, filed on October 28, 1998, each of which is incorporated herein by reference in its entirety.

1. INTRODUCTION

The present invention relates, first, to the *HKNG1* gene, shown herein to be associated with neuropsychiatric disorders, in particular, bipolar affective disorder and schizophrenia. The invention includes recombinant DNA molecules and cloning vectors comprising sequences of the *HKNG1* gene and hosts organisms which have been transformed with such DNA molecules and cloning vectors. The present invention further relates to *HKNG1* gene products, and to antibodies directed against such *HKNG1* gene products. The present invention also relates to methods of using the *HKNG1* gene and gene product, including drug screening assays, and diagnostic and therapeutic methods for the treatment of *HKNG1*-mediated disorders, including *HKNG1*-mediated neuropsychiatric disorders such as bipolar affective disorder.

2. BACKGROUND OF THE INVENTION

There are only a few psychiatric disorders in which clinical manifestations of the disorder can be correlated with demonstrable defects in the structure and/or function of the nervous system. Well-known examples of such disorders include Huntington's disease, which can be traced to a mutation in a single gene and in which neurons in the striatum degenerate, and Parkinson's disease, in which dopaminergic neurons in the nigro-striatal pathway degenerate. The vast majority of psychiatric disorders, however, presumably involve subtle and/or undetectable changes, at the cellular and/or molecular levels, in

nervous system structure and function. This lack of detectable neurological defects distinguishes "neuropsychiatric" disorders, such as schizophrenia, attention deficit disorders, schizoaffective disorder, bipolar affective disorders, or unipolar affective disorder, from neurological disorders, in which anatomical or biochemical pathologies are manifest. Hence, identification of the causative defects and the neuropathologies of neuropsychiatric disorders are needed in order to enable clinicians to evaluate and prescribe appropriate courses of treatment to cure or ameliorate the symptoms of these disorders.

One of the most prevalent and potentially devastating of neuropsychiatric disorders is bipolar affective disorder (BAD), also known as bipolar mood disorder (BP) or manic-depressive illness, which is characterized by episodes of elevated mood (mania) and depression (Goodwin, *et al.*, 1990, *Manic Depressive Illness*, Oxford University Press, New York). The most severe and clinically distinctive forms of BAD are BP-I (severe bipolar affective (mood) disorder), which affects 2-3 million people in the United States, and SAD-M (schizoaffective disorder manic type). They are characterized by at least one full episode of mania, with or without episodes of major depression (defined by lowered mood, or depression, with associated disturbances in rhythmic behaviors such as sleeping, eating, and sexual activity). BP-I often co-segregates in families with more etiologically heterogeneous syndromes, such as with a unipolar affective disorder such as unipolar major depressive disorder (MDD), which is a more broadly defined phenotype (Freimer and Reus, 1992, in *The Molecular and Genetic Basis of Neurological Disease*, Rosenberg, *et al.*, eds., Butterworths, New York, pp. 951-965; McInnes and Freimer, 1995, Curr. Opin. Genet. Develop., 5, 376-381). BP-I and SAD-M are severe mood disorders that are frequently difficult to distinguish from one another on a cross-sectional basis, follow similar clinical courses, and segregate together in family studies (Rosenthal, *et al.*, 1980, Arch. General Psychiat. 37, 804-810; Levinson and Levitt, 1987, Am. J. Psychiat. 144, 415-426; Goodwin, *et al.*, 1990, *Manic Depressive Illness*, Oxford University Press, New York). Hence, methods for distinguishing neuropsychiatric disorders such as these are needed in order to effectively diagnose and treat afflicted individuals.

Currently, individuals are typically evaluated for BAD using the criteria set forth in the most current version of the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM). While many drugs have been used to treat

individuals diagnosed with BAD, including lithium salts, carbamazepine and valproic acid, none of the currently available drugs are adequate. For example, drug treatments are effective in only approximately 60-70% of individuals diagnosed with BP-I. Moreover, it is currently impossible to predict which drug treatments will be effective in, for example, particular BP-I affected individuals. Commonly, upon diagnosis, affected individuals are prescribed one drug after another until one is found to be effective. Early prescription of an effective drug treatment, therefore, is critical for several reasons, including the avoidance of extremely dangerous manic episodes and the risk of progressive deterioration if effective treatments are not found.

The existence of a genetic component for BAD is strongly supported by segregation analyses and twin studies (Bertelson, *et al.*, 1977, Br. J. Psychiat. 130, 330-351; Freimer and Reus, 1992, in *The Molecular and Genetic Basis of Neurological Disease*, Rosenberg, *et al.*, eds., Butterworths, New York, pp. 951-965; Pauls, *et al.*, 1992, Arch. Gen. Psychiat. 49, 703-708). Efforts to identify the chromosomal location of genes that might be involved in BP-I, however, have yielded disappointing results in that reports of linkage between BP-I and markers on chromosomes X and 11 could not be independently replicated nor confirmed in the re-analyses of the original pedigrees, indicating that with BAD linkage studies, even extremely high lod scores at a single locus, can be false positives (Baron, *et al.*, 1987, Nature 326, 289-292; Egeland, *et al.*, 1987, Nature 325, 783-787; Kelsoe, *et al.*, 1989, Nature 342, 238-243; Baron, *et al.*, 1993, Nature Genet. 3, 49-55).

Recent investigations have suggested possible localization of BAD genes on chromosomes 18p and 21q, but in both cases the proposed candidate region is not well defined and no unequivocal support exists for either location (Berrettini, *et al.*, 1994, Proc. Natl. Acad. Sci. USA 91, 5918-5921; Murray, *et al.*, 1994, Science 265, 2049-2054; Pauls, *et al.*, 1995, Am. J. Hum. Genet. 57, 636-643; Maier, *et al.*, 1995, Psych. Res. 59, 7-15; Straub, *et al.*, 1994, Nature Genet. 8, 291-296).

Mapping genes for common diseases believed to be caused by multiple genes, such as BAD, may be complicated by the typically imprecise definition of phenotypes, by etiologic heterogeneity, and by uncertainty about the mode of genetic transmission of the disease trait. With neuropsychiatric disorders there is even greater ambiguity in distinguishing individuals who likely carry an affected genotype from those

who are genetically unaffected. For example, one can define an affected phenotype for BAD by including one or more of the broad grouping of diagnostic classifications that constitute the mood disorders: BP-I, SAD-M, MDD, and bipolar affective (mood) disorder with hypomania and major depression (BP-II).

Thus, one of the greatest difficulties facing psychiatric geneticists is uncertainty regarding the validity of phenotype designations, since clinical diagnoses are based solely on clinical observation and subjective reports. Also, with complex traits such as neuropsychiatric disorders, it is difficult to genetically map the trait-causing genes because: (1) neuropsychiatric disorder phenotypes do not exhibit classic Mendelian recessive or dominant inheritance patterns attributable to a single genetic locus, (2) there may be incomplete penetrance, *i.e.*, individuals who inherit a predisposing allele may not manifest disease; (3) a phenocopy phenomenon may occur, *i.e.*, individuals who do not inherit a predisposing allele may nevertheless develop disease due to environmental or random causes; (4) genetic heterogeneity may exist, in which case mutations in any one of several genes may result in identical phenotypes.

Despite these difficulties, however, identification of the chromosomal location, sequence and function of genes and gene products responsible for causing neuropsychiatric disorders such as bipolar affective disorders is of great importance for genetic counseling, diagnosis and treatment of individuals in affected families.

3. SUMMARY OF THE INVENTION

The present invention relates, first, to the discovery, identification, and characterization of novel nucleic acid molecules that are associated with human neuropsychiatric disorders, such as bipolar affective disorder (BAD) and schizophrenia. The invention further relates to the discovery, identification, and characterization of proteins encoded by such nucleic acid molecules, or by degenerate, *e.g.*, allelic, variants thereof.

In particular, the nucleic acid molecules of the present invention represent nucleic acid sequences corresponding to the gene referred to herein as *HKNG1*. As demonstrated in the Examples presented below in Sections 6 and 8, the *HKNG1* gene is associated with human neuropsychiatric disorders, in particular BAD. The *HKNG1* gene is associated with other human neuropsychiatric disorders as well, such as schizophrenia.

In addition to the positive correlation between mutations within the *HKNG1* gene and individuals exhibiting symptoms of BAD, described in Section 6 and 8, the present invention is further based, in part, on Applicants' discovery of novel *HKNG1* cDNA sequences. Applicants' discovery of such cDNA sequences has led to the elucidation of the *HKNG1* genomic (that is, intron/exon) structure, and to the discovery of full-length and alternately spliced *HKNG1* variants and the polypeptides encoded by such variants. These discoveries are described in Section 7, below. Applicants' discovery of the *NKNG1* gene and its gene product has also led to the identification, described in Section 9 below, of a receptor for HKNG1 -- LRP-2. Further, the receptor associated protein (RAP) is shown in Section 9 to inhibit HKNG1 binding to LRP-2, and genetic linkage data summarized in Section 10 shows that RAP is also associated with neuropsychiatric disorders such as BAD and schizophrenia.

The invention encompasses nucleic acid molecules which comprise the following nucleotide sequences: (a) nucleotide sequences (e.g., SEQ ID NOS: 1, 3, 5, and 6) that encode a *HKNG1* gene product (SEQ ID NO: 2; SEQ ID NO: 4), as well as allelic variants and homologues thereof; (b) nucleotide sequences comprising the novel *HKNG1* sequences disclosed herein that encode mutants of the *HKNG1* gene product in which all or a part of one or more of the *HKNG1* domains is deleted or altered; (c) nucleotide sequences that encode fusion proteins comprising a *HKNG1* gene product (SEQ ID NO: 2; SEQ ID NO: 4), or a portion thereof fused to a heterologous polypeptide; and (d) nucleotide sequences within the *HKNG1* gene, as well as chromosome 18p nucleotide sequences flanking the *HKNG1* gene, which can be utilized as part of the methods of the invention for identifying and diagnosing individuals at risk for or exhibiting an *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia. The nucleic acid molecules of (a) through (d), above, can include, but are not limited to, cDNA and genomic DNA sequences.

The invention also encompasses the expression products of the nucleic acid molecules listed above; *i.e.*, proteins and/or polypeptides that are encoded by the above *HKNG1* nucleic acid molecules, or by degenerate, *e.g.*, allelic, variants thereof.

The compositions of the present invention further encompass agonists and antagonists of the *HKNG1* gene product, including small molecules, large molecules, and antibodies, as well as nucleotide sequences that can be used to inhibit *HKNG1* gene

expression (*e.g.*, antisense and ribozyme molecules, and gene or regulatory sequence replacement constructs) or to enhance *HKNG1* gene expression (*e.g.*, expression constructs that place the *HKNG1* gene under the control of a strong promoter system).

The compositions of the present invention further include cloning vectors, including expression vectors, containing the nucleic acid molecules of the invention, as well as hosts which have been transformed with such nucleic acid molecules. Such hosts include transgenic non-human animals, particularly non-human mammals, that express an *HKNG1* transgene, or "knock-outs" that do not express *HKNG1*.

Transgenic non-human animals of the invention include animals that express an *HKNG1* transgene at higher or lower levels than normal, wild-type animals. The transgenic animals of the invention also include animals that express a mutant variant or polymorphism of an *HKNG1* transgene which is associated with *HKNG1*-mediated neuropsychiatric disorders, such as BAD and schizophrenia.

The invention further relates to methods for the treatment of *HKNG1*-mediated neuropsychiatric disorders, wherein such methods comprise administering a compound which modulates the expression of a mammalian *HKNG1* gene and/or the synthesis or activity of a mammalian *HKNG1* gene product so symptoms of the disorder are ameliorated.

The invention further relates to methods for the treatment of mammalian *HKNG1*-mediated neuropsychiatric disorders resulting from *HKNG1* gene mutations, wherein such methods comprise supplying the mammal with a nucleic acid molecule encoding an unimpaired *HKNG1* gene product such that an unimpaired *HKNG1* gene product is expressed and symptoms of the disorder are ameliorated.

The invention further relates to methods for the treatment of mammalian *HKNG1*-mediated neuropsychiatric disorders resulting from *HKNG1* gene mutations, wherein such methods comprise supplying the mammal with a cell comprising a nucleic acid molecule that encodes an unimpaired *HKNG1* gene product such that the cell expresses the unimpaired *HKNG1* gene product and symptoms of the disorder are ameliorated.

The invention also encompasses pharmaceutical formulations and methods for treating neuropsychiatric disorders, such as BAD and schizophrenia, involving *HKNG1* gene.

In addition, the present invention is directed to methods that utilize the *HKNG1* nucleic acid sequences, chromosome 18p nucleotide sequences flanking the *HKNG1* gene and/or *HKNG1* gene product sequences for the diagnostic evaluation, genetic testing and prognosis of a *HKNG1*-mediated neuropsychiatric disorder. For example, the invention relates to methods for diagnosing *HKNG1*-mediated neuropsychiatric disorders, wherein such methods comprise measuring *HKNG1* gene expression in a patient sample, or detecting a *HKNG1* mutation in the genome of a mammal, including a human, suspected of exhibiting such a disorder. In one embodiment, nucleic acid molecules encoding *HKNG1* can be used as diagnostic hybridization probes or as primers for diagnostic PCR analysis for the identification of *HKNG1* gene mutations, allelic variations and regulatory defects in the *HKNG1* gene which correlate with neuropsychiatric disorders such as BAD or schizophrenia.

The invention still further relates to methods for identifying compounds which modulate the expression of the mammalian *HKNG1* gene and/or the synthesis or activity of the mammalian *HKNG1* gene products, including therapeutic compounds, which reduce or eliminate the symptoms of *HKNG1*-mediated disorders, including *HKNG1*-mediated neuropsychiatric disorders such as BAD and schizophrenia. In particular, cellular and non-cellular assays are described that can be used to identify compounds that interact with the *HKNG1* gene product, e.g., modulate the activity of the *HKNG1* and/or bind to the *HKNG1* gene product. Such cell-based assays of the invention utilize cells, cell lines, or engineered cells or cell lines that express the *HKNG1* gene product.

In one embodiment, such methods comprise contacting a compound to a cell that expresses a *HKNG1* gene, measuring the level of *HKNG1* gene expression, gene product expression or gene product activity, and comparing this level to the level of *HKNG1* gene expression, gene product expression or gene product activity produced by the cell in the absence of the compound, such that if the level obtained in the presence of the compound differs from that obtained in its absence, a compound that modulates the expression of the mammalian *HKNG1* gene and/or the synthesis or activity of the mammalian *HKNG1* gene products has been identified.

In another embodiment, such methods comprise administering a compound to a host, such as a transgenic animal, that expresses an *HKNG1* transgene or a mutant

HKNG1 transgene associated with an *HKNG1*-mediated disorder such as a neuropsychiatric disorder (*e.g.*, BAD or schizophrenia), or to an animal, *e.g.*, a knock-out animal, that does not express *HKNG1*, and measuring the level of *HKNG1* gene expression, gene product expression, gene product activity, or symptoms of an *HKNG1*-mediated disorder such as an *HKNG1*-mediated neuropsychiatric disorder (*e.g.*, BAD or schizophrenia). The measured level is compared to the level obtained in a host that is not exposed to the compound, such that if the level obtained when the host is exposed to the compound differs from that obtained in a host not exposed to the compound, a compound modulates the expression of the mammalian *HKNG1* gene and/or the synthesis or activity of the mammalian *HKNG1* gene products, and/or the symptoms of an *HKNG1*-mediated disorder such as a neuropsychiatric disorder (*e.g.*, BAD or schizophrenia), has been identified.

The present invention still further relates to methods for selecting an effective drug to administer to an individual having a *HKNG1*-medicated disorder. Such methods are based on the detection of genetic polymorphisms in the *HKNG1* gene or variations in *HKNG1* gene expression due to altered methylation, differential splicing, or post-translational modification of the *HKNG1* gene product which can affect the safety and efficacy of a therapeutic agent.

As briefly discussed above, the present invention is based, in part, on the genetic and physical mapping of the *HKNG1* gene to a specific portion of the short arm of human chromosome 18 that is associated with human neuropsychiatric disorders, in particular, bipolar affective disorder. These results are described in the Example presented, below, in Section 6. The invention is also based on the elucidation of the *HKGN1* nucleotide sequence, amino acid sequence and expression pattern, as described in the Example presented, below, in Section 7. The invention is further based on the identification of specific mutations and/or polymorphisms within the *HKNG1* gene which positively correlate with neuropsychiatric disorders, in particular, BAD, as described in the Example presented below in Section 8. These mutations include a point mutation discovered in an individual affected by BAD which is absent from the corresponding wild-type nucleic acid derived from non-affected individuals and linkage analysis of three markers showing cosegregation with a population of individuals with BAD. This mutation is a single base mutation which results in a mutant *HKNG1* gene product comprising substitution of a lysine

residue for the wild-type glutamic acid residue at *HKNG1* amino acid position 202 of the polypeptide of SEQ ID NO:2 or the *HKNG1* amino acid residue 184 of the polypeptide of SEQ ID NO:4.

3.1. DEFINITIONS

As used herein, the following terms shall have the abbreviations indicated.

BAC, bacterial artificial chromosomes
BAD, bipolar affective disorder(s)
BP, bipolar mood disorder
BP-I, severe bipolar affective (mood) disorder
BP-II, bipolar affective (mood) disorder with hypomania and major depression
bp, base pair(s)
EST, expressed sequence tag
lod, logarithm of odds
MDD, unipolar major depressive disorder
ROS, reactive oxygen species
RT-PCR, reverse transcriptase PCR
SSCP, single-stranded conformational polymorphism
SAD-M, schizoaffective disorder manic type
STS, short tag sequence
YAC, yeast artificial chromosome

For convenience, the meaning of certain terms and phrases employed in the specification, examples, and appended claims are provided below.

The term "allele", which is used interchangeably herein with "allelic variant" refers to alternative forms of a gene, nucleic acid or portions thereof, as well as to a polypeptide encoded by said gene, nucleic acid, or portion thereof. Nucleic acid alleles occupy the same locus or position on homologous chromosomes. When a subject has two identical alleles of a gene, the subject is said to be homozygous for the gene or allele. When

a subject has two different alleles of a gene, the subject is said to be heterozygous for the gene. Alleles of a specific gene can differ from each other in a single nucleotide, or several nucleotides, and can include substitutions, deletions, and insertions of nucleotides. An allele of a gene can also be a form of a gene containing a mutation.

The term "allelic variant of a polymorphic region of an *HKNG* gene" refers to a region of an *HKNG* gene having one of several nucleotide sequences found in that region of the gene in other individuals, as well as to polypeptides encoded by nucleic acid molecules comprising said sequences.

The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds an antigen, such as a polypeptide of the invention, *e.g.*, an epitope of a polypeptide of the invention. A molecule which specifically binds to a given polypeptide of the invention is a molecule which binds the polypeptide, but does not substantially bind other molecules in a sample, *e.g.*, a biological sample, which naturally contains the polypeptide. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope.

"Cells," "host cells" or "recombinant host cells" are terms used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

"Complementary" sequences as used herein refer to sequences which have sufficient complementarity to be able to hybridize, forming a stable duplex.

A "delivery complex" shall mean a targeting means (*e.g.*, a molecule that results in higher affinity binding of a gene, protein, polypeptide or peptide to a target cell surface and/or increased cellular uptake by a target cell). Examples of targeting means

include: sterols (*e.g.*, cholesterol), lipids (*e.g.*, a cationic lipid, virosome or liposome), viruses (*e.g.*, adenovirus, adeno-associated virus, and retrovirus) or target cell specific binding agents (*e.g.*, ligands recognized by target cell specific receptors). Preferred complexes are sufficiently stable *in vivo* to prevent significant uncoupling prior to internalization by the target cell. However, the complex is cleavable under appropriate conditions within the cell so that the gene, protein, polypeptide or peptide is released in a functional form. It is also possible that soluble forms of the protein also exist. Such soluble isoforms can arise through variable splicing of the *HKNG* gene or alternatively as a result of proteolysis of a membranous isoform.

As is well known, genes for a particular polypeptide may exist in single or multiple copies within the genome of an individual. Such duplicate genes may be identical or may have certain modifications, including nucleotide substitutions, additions or deletions, which all still code for polypeptides having substantially the same activity. The term "DNA sequence encoding an HKNG polypeptide" may thus refer to one or more genes within a particular individual. Moreover, certain differences in nucleotide sequences may exist between individual organisms, which are called alleles. Such allelic differences may or may not result in differences in amino acid sequence of the encoded polypeptide yet still encode a protein with the same biological activity.

As used herein, the term "gene" or "recombinant gene", as applied to *HKNG*, refers to a polynucleotide or nucleic acid molecule comprising an open reading frame encoding one of the HKNG polypeptides of the present invention. In one embodiment, these terms relate to a cDNA sequence including, but not limited to, a polynucleotide or nucleic acid sequence obtained via reverse transcription of an mRNA molecule. In one embodiment, the term nucleic acid or polynucleotide is a nucleic acid molecule which is not genomic but is a cDNA derived from a contiguous coding region which includes, but is not limited to, reverse transcribed cDNA. In another embodiment, the term nucleic acid or polynucleotide refers to a nucleic acid molecule which comprises contiguous nucleotide codons. In yet another embodiment, the term nucleic acid or polynucleotide is a nucleic acid molecule which is genomic but which excludes intronic sequences.

"Homology" or "identity" or "similarity" refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology can be determined

by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base or amino acid, then the molecules are identical at that position. A degree of homology or similarity or identity between nucleic acid sequences is a function of the number of identical or matching nucleotides at positions shared by the nucleic acid sequences. A degree of identity of amino acid sequences is a function of the number of identical amino acids at positions shared by the amino acid sequences. Likewise, a degree of identity of nucleic acid sequences is a function of the number of identical nucleic acids at positions shared by the nucleic acid sequences.

Furthermore, a degree of homology or similarity of amino acid sequences is a function of the number of conserved amino acids at positions shared by the amino acid sequences. A sequence which is "unrelated" or "non-homologous" with one of the human *HKNG* sequences of the present invention typically is a sequence which shares less than 40 % identity, though preferably less than 25 % identity with one of the human *HKNG* sequences of the present invention.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = # of identical positions/total # of positions (*e.g.*, overlapping positions) x 100).

Preferably, the determination of percent identity between two sequences is accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, et al. (1990) *J. Mol.*

Biol. 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules. *Id.* When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. *See* http://www.ncbi.nlm.nih.gov. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, typically only exact matches are counted.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% (65%, 70%, preferably 75% or more) identical to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 2.0 X SSC at 50° C. (low stringency) or 0.2 X SSC, 0.1% SDS at 50-65°C (high stringency). In one embodiment, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID Nos. 1, 3, 5, 6, or 7, or complement thereof, corresponds to a naturally-occurring nucleic acid molecule.

The *HKNG1*-mediated neuropsychiatric disorders referred to herein include, but are not limited to, schizophrenia disorders as well as bipolar affective disorder, *e.g.*, severe bipolar affective (mood) disorder (BP-I), bipolar affective (mood) disorder with hypomania and major depression (BP-II).

The term "*HKNG1*-mediated neuropsychiatric disorder" as used herein refers to a disorder involving an aberrant level of *HKNG1* gene expression, gene product synthesis and/or gene product activity relative to levels found in normal, unaffected, unimpaired individuals, levels found in clinically normal individuals, and/or levels found in a population whose level represents a baseline, average *HKNG1* level.

4. BRIEF DESCRIPTION OF THE FIGURES

FIG. 1A-1B. The nucleotide sequence of human *HKNG1* cDNA (SEQ ID NO: 1) is depicted on the bottom line. The top line depicts the full length amino acid sequence of human *HKNG1* polypeptide (SEQ ID NO: 2) encoded by the human *HKNG1* cDNA sequence. The nucleotide sequence encoding SEQ ID NO:2 corresponds to SEQ ID NO:5.

FIG. 2A-2B. Nucleotide sequence of an alternately spliced human *HKNG1* variant (SEQ ID NO: 3) is depicted on the bottom line. The derived amino acid sequence of the human *HKNG1* gene product (SEQ ID NO: 4) encoded by this alternately spliced cDNA variant is depicted on to top line. The nucleotide sequence encoding SEQ ID NO:4 corresponds to SEQ ID NO:6

FIG. 3A-3R. Genomic sequence of the human *HKNG1* gene (SEQ ID NO. 7). Exons are in bold and the 3' and 5' UTR's are underlined.

FIG. 4. Summary of in situ hybridization analysis of *HKNG1* mRNA distribution in normal human brain tissue.

FIG. 5. *HKNG1* polymorphisms relative to the *HKNG1* consensus wild type sequence. These polymorphisms were isolated from the NIMH schizophrenia collection.

5. DETAILED DESCRIPTION OF THE INVENTION

5.1. THE *HKNG1* GENE

A human *HKNG1* cDNA sequence (SEQ ID NO: 1) encoding the full length amino acid sequence (SEQ ID NO: 2) of the *HKNG1* polypeptide is shown in FIG. 1A-1B. The human *HKNG1* gene encodes a putatively secreted polypeptide of 495 amino acid residues, as shown in FIG. 1A-1B, and SEQ ID NO: 2. The nucleotide sequence of the portion of the cDNA corresponding to the coding sequence for *HKNG1* (SEQ ID NO:2) is depicted as SEQ ID NO:5 An alternately spliced human *HKNG1* cDNA sequence (SEQ ID NO: 3) encoding a human *HKNG1* variant gene product is shown in FIG. 2A-2B. Thus, the human *HKNG1* gene also encodes a putatively secreted polypeptide of 477 amino acid residues, as shown in FIG. 2A-2B, and SEQ ID NO:4. The nucleotide sequence of the portion of the cDNA corresponding to the coding sequence for *HKNG1* (SEQ ID NO:4) is depicted in SEQ ID NO:6.

The genomic structure of the *HKNG1* gene has been elucidated and is depicted in FIG. 3A-3R, with the eleven *HKNG1* exons indicated in bold type, the 5'- and 3'-untranslated regions indicated by underlining. The consensus wild type genomic sequence of the *HKNG1* gene is depicted in FIG. 3A-3R and SEQ ID NO:7.

The *HKNG1* gene nucleic acid molecules of the invention include: (a) nucleotide sequences (e.g., SEQ ID NOS: 1, 3, 5, and 6) that encode a *HKNG1* gene product (SEQ ID NO: 2; SEQ ID NO: 4), as well as homologues thereof; (b) nucleotide sequences comprising the novel *HKNG1* sequences disclosed herein that encode mutants of the *HKNG1* gene product in which all or a part of one or more of the domains is deleted or altered; (c) nucleotide sequences that encode fusion proteins comprising a *HKNG1* gene product (SEQ ID NO: 2; SEQ ID NO: 4), or a portion thereof fused to a heterologous polypeptide; and (d) nucleotide sequences within the *HKNG1* gene, and chromosome 18p nucleotide sequences flanking the *HKNG1* gene which can be utilized as part of the methods of the invention for identifying and diagnosing individuals at risk for or exhibiting an *HKNG1*-mediated neuropsychiatric disorder, such as BAD.

The *HKNG1* nucleotide sequences of the invention further include: (a) any nucleotide sequence that hybridizes to the complement of a nucleic acid molecule that encodes an *HKNG1* gene product under highly stringent conditions, e.g., hybridization to

filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, at p. 2.10.3). Preferentially, such nucleic acid molecules encode gene products functionally equivalent to an *HKNG1* gene product; and (b) any nucleotide sequence that hybridizes to the complement of a nucleic acid molecule that encodes an *HKNG1* gene product under less stringent conditions, such as moderately stringent conditions, e.g., washing in 0.2xSSC/0.1% SDS at 42°C (Ausubel et al., 1989, supra), yet which still encodes a functionally equivalent *HKNG1* gene product. Functional equivalents of *HKNG1* include naturally occurring *HKNG1* present in the same or different species. In one embodiment, *HKNG1* gene sequences in non-human species map to chromosome regions synthetic to the human 18p chromosome location within which human *HKNG1* lies.

The invention also includes nucleic acid molecules, preferably DNA molecules, that are the complements of the nucleotide sequences of the preceding paragraphs. Among the nucleic acid molecules of the invention are deoxyoligonucleotides ("oligos") which hybridize under highly or moderately stringent conditions to the *HKNG1* nucleic acid molecules described above. Examplary highly stringent conditions may refer, e.g., to washing in 6xSSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos). These nucleic acid molecules may encode or act as antisense molecules, useful, for example, in *HKNG1* gene regulation, and/or as antisense primers in amplification reactions of *HKNG1* gene nucleic acid sequences. Further, such sequences may be used as part of ribozyme and/or triple helix sequences, also useful for *HKNG1* gene regulation. Still further, such molecules may be used as components of diagnostic methods whereby, for example, the presence of a particular *HKNG1* allele involved in a neuropsychiatric disorder, such as BAD, may be detected.

The *HKNG1* nucleotide sequences of the invention can be readily obtained, for example, by standard sequencing and the sequence provided herein.

With respect to *HKNG1* allelic variants, such variants include, but are not limited to, the variants summarized in FIG. 5 and a variant which encodes a full length *HKNG1* polypeptide comprising a substitution of a lysine amino acid at amino acid residue

202 of the *HKNG1* polypeptide shown in FIG. 1A-1B and SEQ ID NO:2 or the *HKNG1* amino acid residue 184 of the polypeptide of SEQ ID NO:4.

With respect to the cloning of allelic variants of the human *HKNG1* gene and homologues from other species (*e.g.*, mouse), the isolated *HKNG1* gene sequences disclosed herein may be labeled and used to screen a cDNA library constructed from mRNA obtained from appropriate cells or tissues (*e.g.*, brain tissues) derived from the organism (*e.g.*, mouse) of interest. The hybridization conditions used should be of a lower stringency when the cDNA library is derived from an organism different from the type of organism from which the labeled sequence was derived.

Alternatively, the labeled fragment may be used to screen a genomic library derived from the organism of interest, again, using appropriately stringent conditions. Low stringency conditions are well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook, *et al.*, 1989, Molecular Cloning, A Laboratory Manual, Second Edition, Cold Spring Harbor Press, N.Y.; and Ausubel, *et al.*, 1989, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.

Further, a *HKNG1* gene allelic variant may be isolated from, for example, human nucleic acid, by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of amino acid sequences within the *HKNG1* gene product disclosed herein. The template for the reaction may be cDNA obtained by reverse transcription of mRNA prepared from, for example, human or non-human cell lines or tissue known or suspected to express a *HKNG1* gene allele (such as, for example, brain cells from individuals having BAD. Preferably, the allelic variant will be isolated from an individual who has a *HKNG1*-mediated neuropsychiatric disorder. One such variant is described in the examples below.

The PCR product may be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a *HKNG1* gene nucleic acid sequence. The PCR fragment may then be used to isolate a full length cDNA clone by a variety of methods. For example, the amplified fragment may be labeled and used to screen a

bacteriophage cDNA library. Alternatively, the labeled fragment may be used to isolate genomic clones via the screening of a genomic library.

PCR technology may also be utilized to isolate full length cDNA sequences. For example, RNA may be isolated, following standard procedures, from an appropriate cellular or tissue source (*i.e.*, one known, or suspected, to express the *HKNG1* gene, such as, for example, brain tissue samples obtained through biopsy or post-mortem). A reverse transcription reaction may be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. Thus, cDNA sequences upstream of the amplified fragment may easily be isolated. For a review of cloning strategies that may be used, see *e.g.*, Sambrook *et al.*, 1989, *supra*.

A cDNA of a mutant allelic variant of the *HKNG1* gene may be isolated, for example, by using PCR, a technique that is well known to those of skill in the art. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual putatively carrying the mutant *HKNG1* allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant *HKNG1* allele to that of the normal *HKNG1* allele, the mutation(s) responsible for the loss or alteration of function of the mutant *HKNG1* gene product can be ascertained.

Alternatively, a genomic library can be constructed using DNA obtained from an individual suspected of or known to carry a mutant *HKNG1* allele, or a cDNA library can be constructed using RNA from a tissue known, or suspected, to express a mutant *HKNG1* allele. An unimpaired *HKNG1* gene or any suitable fragment thereof may then be labeled and used as a probe to identify the corresponding mutant *HKNG1* allele in such libraries.

Clones containing the mutant *HKNG1* gene sequences may then be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

Additionally, an expression library can be constructed utilizing cDNA synthesized from, for example, RNA isolated from a tissue known, or suspected, to express a mutant *HKNG1* allele in an individual suspected of or known to carry such a mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against the normal *HKNG1* gene product, as described, below, in Section 5.3. (For screening techniques, see, for example, Harlow and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor.)

In cases where a *HKNG1* mutation results in an expressed gene product with altered function (*e.g.*, as a result of a missense or a frameshift mutation), a polyclonal set of anti-*HKNG1* gene product antibodies are likely to cross-react with the mutant *HKNG1* gene product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

HKNG1 mutations can further be detected using PCR amplification techniques. Primers can routinely be designed to amplify overlapping regions of the whole *HKNG1* sequence including the promoter regulating region. In one embodiment, primers are designed to cover the exon-intron boundaries such that, coding regions can be scanned for mutations. Exemplary primers for analyzing *HKNG1* exons are provided in Table I, of Section 5.6, below.

5.2. PROTEIN PRODUCTS OF THE *HKNG1* GENE

HKNG1 gene products, or peptide fragments thereof, can be prepared for a variety of uses. For example, such gene products, or peptide fragments thereof, can be used for the generation of antibodies, in diagnostic assays, or for the identification of other cellular or extracellular gene products involved in the regulation of *HKNG1*-mediated neuropsychiatric disorders, such as BAD.

The amino acid sequences depicted in FIG. 1A-1B and FIG. 2A-2B represent *HKNG1* gene products. *HKNG1* gene product, sometimes referred to herein as an "*HKNG1*

protein", includes those gene products encoded by the *HKNG1* gene sequences depicted in FIG. 1A-1B and FIG. 2A-2B, as well as other human allelic variants or non-human variants of *HKNG1* that can be identified by the methods herein described.

In addition, *HKNG1* gene products may include proteins that represent functionally equivalent gene products. Functionally equivalent gene products may include, for example, gene products encoded by one of the *HKNG1* nucleic acid molecules described in Section 5.1, above. In preferred embodiments, such functionally equivalent *HKNG1* gene products are naturally occuring gene products. Such an equivalent *HKNG1* gene product may contain deletions, including internal deletions, additions, including additions yielding fusion proteins, or substitutions of amino acid residues within and/or adjacent to the amino acid sequence encoded by the *HKNG1* gene sequences described, above, in Section 5.1, but that result in a "silent" change, in that the change produces a *HKNG1* gene product with the same activity as the *HKNG1* gene product depicted in FIG. 1A-1B or FIG. 2A-2B.

Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

Alternatively, where alteration of function is desired, deletion or non-conservative alterations can produce altered, including reduced-activity, *HKNG1* gene products. Such alterations can, for example, alter one or more of the biological functions of the *HKNG1* gene product. Further, such alterations can be selected so as to generate *HKNG1* gene products that are better suited for expression, scale up, etc. in the host cells chosen. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

As another example, altered *HKNG1* gene products can be engineered that correspond to variants of the *HKNG1* gene product associated with *HKNG1*-mediated neuropsychiatric disorders such as BAD. Such altered *HKNG1* gene products include, but are not limited to, *HKNG1* proteins or peptides comprising substitution of a lysine residue

for the wild-type glutamic acid residue at *HKNG1* amino acid position 202 in FIG. 1A-1B (SEQ ID NO:2) or amino acid position 184 (SEQ ID NO:4) in FIG. 2A-2B.

Peptides and/or proteins corresponding to one or more domains of the *HKNG1* protein as well as fusion proteins and/or in which the full length *HKNG1* protein, an *HKNG1* peptide, or a truncated HKNG1 protein or peptide is fused to an unrelated protein are also within the scope of this invention. Such proteins and peptides can be designed on the basis of the *HKNG1* nucleotide sequence disclosed in Section 5.1, above, and/or on the basis of the *HKNG1* amino acid sequence disclosed in the Section. Such fusion proteins include, but are not limited to, IgFc fusions which stabilize the *HKNG1* protein or peptide and prolong half life *in vivo*; or fusions to any amino acid sequence that allows the fusion protein to be anchored to the cell membrane; or fusions to an enzyme, fluorescent protein, or luminescent protein which provides a marker function.

The *HKNG1* gene products, peptide fragments thereof and fusion proteins thereof, may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the *HKNG1* gene products, polypeptides, peptides, fusion peptide and fusion polypeptides of the invention by expressing nucleic acid containing *HKNG1* gene sequences are described herein. Methods that are well known to those skilled in the art can be used to construct expression vectors containing *HKNG1* gene product coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. See, for example, the techniques described in Sambrook, *et al.*, 1989, *supra*, and Ausubel, *et al.*, 1989, *supra*. Alternatively, RNA capable of encoding *HKNG1* gene product sequences may be chemically synthesized using, for example, synthesizers. See, for example, the techniques described in "Oligonucleotide Synthesis", 1984, Gait, ed., IRL Press, Oxford.

A variety of host-expression vector systems may be utilized to express the *HKNG1* gene product coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells that may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the *HKNG1* gene product of the invention *in situ*. These include but are not limited to microorganisms such as bacteria

(*e.g.*, *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing *HKNG1* gene product coding sequences; yeast (*e.g.*, *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing the *HKNG1* gene product coding sequences; insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus) containing the *HKNG1* gene product coding sequences; plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing *HKNG1* gene product coding sequences; or mammalian cell systems (*e.g.*, COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the *HKNG1* gene product being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions of *HKNG1* gene product or for raising antibodies to *HKNG1* gene product, for example, vectors that direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther *et al.*, 1983, EMBO J. 2:1791), in which the *HKNG1* gene product coding sequence may be ligated individually into the vector in frame with the *lacZ* coding region so that a fusion protein is produced; pIN vectors (Inouye and Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke and Schuster, 1989, J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica*, nuclear polyhidrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The *HKNG1* gene product coding sequence may be cloned individually

into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of *HKNG1* gene product coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed. (*e.g.*, see Smith, *et al.*, 1983, J. Virol. 46:584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the *HKNG1* gene product coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing *HKNG1* gene product in infected hosts. (*e.g.*, See Logan and Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted *HKNG1* gene product coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire *HKNG1* gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the *HKNG1* gene coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner, *et al.*, 1987, Methods in Enzymol. 153:516-544).

In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (*e.g.*, glycosylation) and processing (*e.g.*, cleavage) of protein

products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, and WI38.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express the *HKNG1* gene product may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (*e.g.*, promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci that in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines that express the *HKNG1* gene product. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the *HKNG1* gene product.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, *et al.*, 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska and Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine phosphoribosyltransferase (Lowy, *et al.*, 1980, Cell 22:817) genes can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler, *et al.*, 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, *et al.*, 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan and Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo,

which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, *et al.*, 1981, J. Mol. Biol. 150:1); and hygro, which confers resistance to hygromycin (Santerre, *et al.*, 1984, Gene 30:147).

Alternatively, the expression characteristics of an endogenous *HKNG1* gene within a cell line or microorganism may be modified by inserting a heterologous DNA regulatory element into the genome of a stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous *HKNG1* gene. For example, an endogenous *HKNG1* gene which is normally "transcriptionally silent", *i.e.*, an *HKNG1* gene which is normally not expressed, or is expressed only at very low levels in a cell line or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, a transcriptionally silent, endogenous *HKNG1* gene may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such that it is operatively linked with an endogenous *HKNG1* gene, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described *e.g.*, in Chappel, U.S. Patent No. 5,272,071; PCT publication No. WO 91/06667, published May 16, 1991.

Alternatively, any fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht, *et al.* allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, *et al.*, 1991, Proc. Natl. Acad. Sci. USA 88:8972-8976). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni^{2+}·nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

The *HKNG1* gene products can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, guinea pigs, pigs, micro-pigs, goats, sheep, and non-human primates, *e.g.*, baboons, monkeys, and chimpanzees may be used to generate *HKNG1* transgenic animals. The term "transgenic,"

as used herein, refers to animals expressing *HKNG1* gene sequences from a different species (*e.g.*, mice expressing huma H*KNG1* gene sequences), as well as animals that have been genetically engineered to overexpress endogenous (*i.e.*, same species) *HKNG1* sequences or animals that have been genetically engineered to no longer express endogenous *HKNG1* gene sequences (*i.e.*, "knock-out" animals), and their progeny.

Any technique known in the art may be used to introduce a *HKNG1* gene transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to pronuclear microinjection (Hoppe and Wagner, 1989, U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ lines (Van der Putten, *et al.*, 1985, Proc. Natl. Acad. Sci., USA 82:6148-6152); gene targeting in embryonic stem cells (Thompson, *et al.*, 1989, Cell 56:313-321); electroporation of embryos (Lo, 1983, Mol. Cell. Biol. 3:1803-1814); and sperm-mediated gene transfer (Lavitrano *et al.*, 1989, Cell 57:717-723) (For a review of such techniques, see Gordon, 1989, Transgenic Animals, Intl. Rev. Cytol. 115, 171-229)

Any technique known in the art may be used to produce transgenic animal clones containing a *HKNG1* transgene, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal or adult cells induced to quiescence (Campbell, *et al.*, 1996, Nature 380:64-66; Wilmut, *et al.*, Nature 385:810-813).

The present invention provides for transgenic animals that carry a *HKNG1* transgene in all their cells, as well as animals that carry the transgene in some, but not all their cells, *i.e.*, mosaic animals. The transgene may be integrated as a single transgene or in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko *et al.* (Lasko, *et al.*, 1992, Proc. Natl. Acad. Sci. USA 89:6232-6236). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the *HKNG1* transgene be integrated into the chromosomal site of the endogenous *HKNG1* gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous *HKNG1* gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the

nucleotide sequence of the endogenous *HKNG1* gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous *HKNG1* gene in only that cell type, by following, for example, the teaching of Gu, *et al.* (Gu, *et al.*, 1994, Science 265, 103-106). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant *HKNG1* gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to assay whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques that include but are not limited to Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and RT-PCR (reverse transcriptase PCR). Samples of *HKNG1* gene-expressing tissue, may also be evaluated immunocytochemically using antibodies specific for the *HKNG1* transgene product.

HKNG1 proteins can be used to treat neuropsychiatric disorders. Such *HKNG1* gene products include but are not limited to soluble derivatives such as peptides or polypeptides corresponding to one or more domains of the *HKNG1* gene product, particularly *HKNG1* gene products, that are modified such that they are deleted for one or more hydrophobic domains. Alternatively, antibodies to the *HKNG1* protein or anti-idiotypic antibodies that mimic the *HKNG1* gene product (including Fab fragments), antagonists or agonists can be used to treat neuropsychiatric disorders involving *HKNG1*. In yet another approach, nucleotide constructs encoding such *HKNG1* gene products can be used to genetically engineer host cells to express such *HKNG1* gene products in vivo; these genetically engineered cells can function as "bioreactors" in the body delivering a continuous supply of *HKNG1* gene product, *HKNG1* peptides, soluble *HKNG1* polypeptides.

5.3. ANTIBODIES TO *HKNG1* GENE PRODUCTS

Described herein are methods for the production of antibodies capable of specifically recognizing one or more *HKNG1* gene product epitopes or epitopes of

conserved variants or peptide fragments of the *HKNG1* gene products. Further, antibodies that specifically recognize mutant forms of *HKNG1*, are encompassed by the invention.

Such antibodies may include, but are not limited to, polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, $F(ab')_2$ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such antibodies may be used, for example, in the detection of a *HKNG1* gene product in an biological sample and may, therefore, be utilized as part of a diagnostic or prognostic technique whereby patients may be tested for abnormal levels of *HKNG1* gene products, and/or for the presence of abnormal forms of such gene products. Such antibodies may also be utilized in conjunction with, for example, compound screening schemes, as described, below, in Section 5.8, for the evaluation of the effect of test compounds on *HKNG1* gene product levels and/or activity. Additionally, such antibodies can be used in conjunction with the gene therapy techniques described, below, in Section 5.9.2 to, for example, evaluate the normal and/or engineered *HKNG1*-expressing cells prior to their introduction into the patient.

Anti-*HKNG1* gene product antibodies may additionally be used in methods for inhibiting abnormal *HKNG1* gene product activity. Thus, such antibodies may, therefore, be utilized as part of treatment methods for a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia.

For the production of antibodies against a *HKNG1* gene product, various host animals may be immunized by injection with a *HKNG1* gene product, or a portion thereof. Such host animals may include, but are not limited to rabbits, mice, and rats, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*.

Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as a *HKNG1* gene

product, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, host animals such as those described above, may be immunized by injection with *HKNG1* gene product supplemented with adjuvants as also described above.

Polyclonal antibodies can be prepared as described above by immunizing a suitable subject with a polypeptide of the invention as an immunogen. Preferred polyclonal antibody compositions are ones that have been selected for antibodies directed against a polypeptide or polypeptides of the invention. Particularly preferred polyclonal antibody preparations are ones that contain only antibodies directed against a polypeptide or polypeptides of the invention. Particularly preferred immunogen compositions are those that contain no other human proteins such as, for example, immunogen compositions made using a non-human host cell for recombinant expression of a polypeptide of the invention. In such a manner, the only human epitope or epitopes recognized by the resulting antibody compositions raised against this immunogen will be present as part of a polypeptide or polypeptides of the invention.

The antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized polypeptide. If desired, the antibody molecules can be isolated from the mammal (e.g., from the blood) and further purified by well-known techniques, such as protein A chromatography to obtain the IgG fraction. Alternatively, antibodies specific for a protein or polypeptide of the invention can be selected for (*e.g.*, partially purified) or purified by, *e.g.*, affinity chromatography. For example, a recombinantly expressed and purified (or partially purified) protein of the invention is produced as described herein, and covalently or non-covalently coupled to a solid support such as, for example, a chromatography column. The column can then be used to affinity purify antibodies specific for the proteins of the invention from a sample containing antibodies directed against a large number of different epitopes, thereby generating a substantially purified antibody composition, *i.e.*, one that is substantially free of contaminating antibodies. By a substantially purified antibody composition is meant, in this context, that the antibody sample contains at most only 30% (by dry weight) of contaminating antibodies directed against epitopes other than those on the desired protein or polypeptide of the invention, and preferably at most 20%, yet more preferably at most 10%, and most preferably at most 5%

(by dry weight) of the sample is contaminating antibodies. A purified antibody composition means that at least 99% of the antibodies in the composition are directed against the desired protein or polypeptide of the invention.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, may be obtained by any technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique of Kohler and Milstein, (1975, Nature 256:495-497; and U.S. Patent No. 4,376,110), the human B-cell hybridoma technique (Kosbor *et al.*, 1983, Immunology Today 4:72; Cole *et al.*, 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole *et al.*, 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated *in vitro* or *in vivo*. Production of high titers of mAbs *in vivo* makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison, *et al.*, 1984, Proc. Natl. Acad. Sci., 81:6851-6855; Neuberger, *et al.*, 1984, Nature 312:604-608; Takeda, *et al.*, 1985, Nature, 314:452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. (See, *e.g.*, Cabilly et al., U.S. Patent No. 4,816,567; and Boss et al., U.S. Patent No. 4,816397, which are incorporated herein by reference in their entirety.)

In addition, techniques have been developed for the production of humanized antibodies. (See, *e.g.*, Queen, U.S. Patent No. 5,585,089, which is incorporated herein by reference in its entirety.) An immunoglobulin light or heavy chain variable region consists of a "framework" region interrupted by three hypervariable regions, referred to as complementarily determining regions (CDRs). The extent of the framework region and CDRs have been precisely defined (see, "Sequences of Proteins of Immunological Interest", Kabat, E. et al., U.S. Department of Health and Human Services (1983)). Briefly,

humanized antibodies are antibody molecules from non-human species having one or more CDRs from the non-human species and a framework region from a human immunoglobulin molecule.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778; Bird, 1988, Science 242:423-426; Huston, *et al.*, 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward, *et al.*, 1989, Nature 334:544-546) can be adapted to produce single chain antibodies against *HKNG1* gene products. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments that recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragments, which can be produced by pepsin digestion of the antibody molecule and the Fab fragments, which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed (Huse, *et al.*, 1989, Science 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

5.4. USES OF *HKNG1* GENE SEQUENCES GENE PRODUCTS, AND ANTIBODIES

Described herein are various applications of *HKNG1* gene sequences, *HKNG1* gene products, including peptide fragments and fusion proteins thereof, and of antibodies directed against *HKNG1* gene products and peptide fragments thereof. Such applications include, for example, prognostic and diagnostic evaluation of a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia, and the identification of subjects with a predisposition to such disorders, as described, below, in Section 5.5. Additionally, such applications include methods for the treatment of a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia, as described, below, in Section 5.9, and for the identification of compounds that modulate the expression of the *HKNG1* gene and/or the synthesis or activity of the *HKNG1* gene product, as described below, in Section 5.8. Such compounds can include, for example, other cellular products that are involved in mood regulation and in *HKNG1*-mediated neuropsychiatric disorders, such as

BAD or schizophrenia. These compounds can be used, for example, in the amelioration of *HKNG1*-mediated neuropsychiatric disorders, such as BAD and schizophrenia.

5.5. DIAGNOSIS OF *HKNG1*-MEDIATED NEUROPSYCHIATRIC DISORDERS

A variety of methods can be employed for the diagnostic and prognostic evaluation of *HKNG1*-mediated neuropsychiatric disorders and for the identification of subjects having a predisposition to such disorders.

Such methods may, for example, utilize reagents such as the *HKNG1* gene nucleotide sequences described in Sections 5.1, and antibodies directed against *HKNG1* gene products, including peptide fragments thereof, as described, above, in Section 5.3. Specifically, such reagents may be used, for example, for:

(1) the detection of the presence of *HKNG1* gene mutations, or the detection of either over- or under-expression of *HKNG1* gene relative to wild-type *HKNG1* levels of expression;

(2) the detection of over- or under-abundance of *HKNG1* gene product relative to wild-type abundance of *HKNG1* gene product; and

(3) the detection of an aberrant level of *HKNG1* gene product activity relative to wild-type *HKNG1* gene product activity levels.

HKNG1 gene nucleotide sequences can, for example, be used to diagnose a *HKNG1*-mediated neuropsychiatric disorder using, for example, the techniques for *HKNG1* mutation detection described above in Section 5.1.

Mutations at a number of different genetic loci may lead to phenotypes related to neuropsychiatric disorders. Ideally, the treatment of patients suffering from such neuropsychiatric disorder will be designed to target the particular genetic loci containing the mutation mediating the disorder. Genetic polymorphisms have been linked to differences in drug effectiveness. Thus, identification of alterations in the *HKNG1* gene or protein can be utilized in pharmacogenetic methods to optimize therapeutic drug treatments.

In an embodiment of the present invention, polymorphisms in the *HKNG1* gene sequence, or variations in *HKNG1* gene expression due to altered methylation, differential splicing, or post-translational modification of the *HKNG1* gene product, may be

utilized to identify an individual having a disease or condition resulting from a *HKNG1*-mediated disorder and thus define the most effective and safest drug treatment. Assays such as those described herein may be used to identify such polymorphisms or variations in *HKNG1* gene expression activity. Once a polymorphism in the *HKNG1* gene, or a variation in *HKNG1* gene expression has been identified in an individual, an appropriate drug treatment can be prescribed to the individual.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one specific *HKNG1* gene nucleic acid or anti-*HKNG1* gene product antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings, to diagnose patients exhibiting abnormalities of a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia.

For the detection of *HKNG1* gene mutations, any nucleated cell can be used as a starting source for genomic nucleic acid. For the detection of *HKNG1* gene expression or *HKNG1* gene products, any cell type or tissue in which the *HKNG1* gene is expressed may be utilized.

Nucleic acid-based detection techniques are described, below, in Section 5.6. Peptide detection techniques are described, below, in Section 5.7.

5.6. DETECTION OF *HKNG1* NUCLEIC ACID MOLECULES

A variety of methods can be employed to screen for the presence of *HKNG1* gene-specific mutations and to detect and/or assay levels of *HKNG1* nucleic acid sequences.

Mutations within the *HKNG1* gene can be detected by utilizing a number of techniques. Nucleic acid from any nucleated cell can be used as the starting point for such assay techniques, and may be isolated according to standard nucleic acid preparation procedures that are well known to those of skill in the art.

HKNG1 nucleic acid sequences may be used in hybridization or amplification assays of biological samples to detect abnormalities involving *HKNG1* gene structure, including point mutations, insertions, deletions, inversions, translocations and chromosomal rearrangements. Such assays may include, but are not limited to, Southern analyses, single-stranded conformational polymorphism analyses (SSCP), and PCR analyses.

Diagnostic methods for the detection of *HKNG1* gene-specific mutations can involve for example, contacting and incubating nucleic acids obtained from a sample, *e.g.*, derived from a patient sample or other appropriate cellular source with one or more labeled nucleic acid reagents including recombinant DNA molecules, cloned genes or degenerate variants thereof, such as described in Section 5.1, above, under conditions favorable for the specific annealing of these reagents to their complementary sequences within or flanking the *HKNG1* gene. As used herein, the term "patient sample , biological sample or appropriate cellular source" refers to a sample of tissue or fluid suspected of containing a mutated or non-mutated *HKNG1* polynucleotide or polypeptide from an individual including, but not limited to, e.g., blood, plasma, serum, ascites, pleural effusion, thoracentisis, spinal fluid, lymph fluid, bone marrow, the external sections of the skin, respiratory, intestinal, and genito-urinary tracts, stool, urine, sputum, tears, saliva, blood cells, tumors, organs, tissue and samples of in vitro cell culture constituents. In the analysis of whether an *HKNG1* gene contains a mutation, most simply, blood can be drawn and DNA extracted from the cells of the blood. In addition, prenatal diagnosis can be accomplished by testing fetal cells, placental cells or amniotic cells for mutations of the *HKNG1* gene. Alteration of a wild-type *HKNG1* allele, whether, for example, by point mutation or deletion, can be detected by any of the means discussed herein. When the probes are used to detect the presence of the target sequences (for example, in screening for susceptibility to a neuropsychiatric disorder, including for example, without limitation, schizophrenia, attention deficit disorder, a schizoaffective disorder, a bipolar affective disorder or a unipolar affective disorder), the biological sample to be analyzed, such as blood, plasma, serum, ascites, pleural effusion, thoracentisis, spinal fluid, lymph fluid, bone marrow, the external sections of the skin, respiratory, intestinal, and genito-urinary tracts, stool, urine, sputum, tears, saliva, blood cells, tumors, organs, tissue and samples of in vitro cell culture constituents, may be treated, if desired, to extract the nucleic acids. The sample nucleic acid may then be prepared in various ways to facilitate detection of the target sequence; e.g. denaturation, restriction digestion, electrophoresis or dot blotting. The targeted region of the *HKNG1* nucleic acid usually must be at least partially single-stranded to form hybrids with the targeting sequence of the probe. If the sequence is naturally single-stranded, denaturation will not be required.

However, if the sequence is double-stranded, the sequence will probably need to be denatured. Denaturation can be carried out by various techniques known in the art.

The diagnostic methods of the present invention further encompass contacting and incubating nucleic acids for the detection of single nucleotide mutations or polymorphisms of the *HKNG1* gene. Preferably, these nucleic acid reagents sequences within the *HKNG1* gene, or chromosome 18p nucleotide sequences flanking the *HKNG1* gene that are 15 to 30 nucleotides in length.

After incubation, all non-annealed nucleic acids are removed from the nucleic acid: *HKNG1* molecule hybrid. The presence of nucleic acids that have hybridized, if any such molecules exist, is then detected. Using such a detection scheme, the nucleic acid from the cell type or tissue of interest can be immobilized, for example, to a solid support such as a membrane, or a plastic surface such as that on a microtiter plate or polystyrene beads. In this case, after incubation, non-annealed, labeled nucleic acid reagents of the type described in Section 5.1 are easily removed. Detection of the remaining, annealed, labeled *HKNG1* nucleic acid reagents is accomplished using standard techniques well-known to those in the art. The *HKNG1* gene sequences to which the nucleic acid reagents have annealed can be compared to the annealing pattern expected from a normal *HKNG1* gene sequence in order to determine whether a *HKNG1* gene mutation is present.

In a preferred embodiment, *HKNG1* mutations or polymorphisms can be detected by using a microassay of *HKNG1* nucleic acid sequences immobilized to a substrate or "gene chip" (see, *e.g.* Cronin, et al., 1996, Human Mutation 7:244-255).

Alternative diagnostic methods for the detection of *HKNG1* gene specific nucleic acid molecules, in patient samples or other appropriate cell sources, may involve their amplification, *e.g.*, by PCR (the experimental embodiment set forth in Mullis, 1987, U.S. Patent No. 4,683,202), followed by the analysis of the amplified molecules using techniques well known to those of skill in the art, such as, for example, those listed above. The resulting amplified sequences can be compared to those that would be expected if the nucleic acid being amplified contained only normal copies of the *HKNG1* gene in order to determine whether an *HKNG1* gene mutation exists.

Among those *HKNG1* nucleic acid sequences which are preferred for such amplification-related diagnostic screening analyses are oligonucleotide primers which

amplify *HKNG1* exon sequences. The sequences of such oligonucleotide primers are, therefore, preferably derived from *HKNG1* intron sequences so that the entire exon, or coding region, can be analyzed as discussed below. Primer pairs useful for amplification of *HKNG1* exons are preferably derived from adjacent introns. Appropriate primer pairs can be chosen such that each of the eleven *HKNG1* exons are amplified. Primers for the amplification of *HKNG1* exons can be routinely designed by one of ordinary skill in the art by utilizing the exon and intron sequences of *HKNG1* shown in Figure 3A-3R.

As an example, and not by way of limitation, Table 1, below, lists primers and primer pairs which can be utilized for the amplification of each of the human *HKGN1* exons one through eleven. In this table, a primer pair is listed for each exon which consists of a forward primer derived from intron sequence upstream of the exon to be amplified, and a reverse primer derived from intron sequence downstream of the exon to be amplified. For exons greater than about 300 base pairs in length, i.e., exons 4 and 7, two primer pairs are listed (marked 4a, 4b, 7a and 7b). Each of the primer pairs can be utilized, therefore, as part of a standard PCR reaction to amplify an individual *HKNG1* exon (or portion thereof). Primer sequences are depicted in a 5' to 3' orientation.

TABLE 1

	Primer Sequence		
1	cggggttggtttccacc	(SEQ ID NO:8)	forward
	gcgaggagagaaatctggg	(SEQ ID NO:9)	reverse
2	tgctcactactttgcagtgttc	(SEQ ID NO:10)	forward
	tgagatcgtgtcactgcattct	(SEQ ID NO:11)	reverse
3	gtaaatctcaaaatgttgggttaatag	(SEQ ID NO:12)	forward
	ctaactcttcttctatcattactc	(SEQ ID NO:13)	reverse
4A	tgtttattgtgtgtctgctgtg	(SEQ ID NO:14)	forward
	ggacaaccaacatgcaaacag	(SEQ ID NO:15)	reverse

	Primer Sequence		
4B	cccaggtgttttcaattgatgc	(SEQ ID NO:16)	foward
	agcagttttgtccttccaagtg	(SEQ ID NO:17)	reverse
5	gtgttttgtaatctgatcagatctc	(SEQ ID NO:18)	forward
	gcagtatttctggtccagatc	(SEQ ID NO:19)	reverse
6	ggtgcacatagatcatgaaatgg	(SEQ ID NO:20)	forward
	taagctgaaataggtgccttaag	(SEQ ID NO:21)	reverse
7A	tttattccatttctgtcccctac	(SEQ ID NO:22)	forward
	aaggctcagttaggtctgtatc	(SEQ ID NO:23)	reverse
7B	caggagttttaacgtcttcagac	(SEQ ID NO:24)	forward
	gactcagaaatgtctaccatttc	(SEQ ID NO:25)	reverse
8	tgtctccacttcttcaaagtgc	(SEQ ID NO:26)	forward
	caaaatgtacctgagaacttaaag	(SEQ ID NO:27)	reverse
9	cacctccaagtttcatggac	(SEQ ID NO:28)	forward
	caaggtatgcacgtgtcatttc	(SEQ ID NO:29)	reverse
10	gaatgtgtattgggatttagtaaac	(SEQ ID NO:30)	forward
	ttgagaattaactattcctgtcaac	(SEQ ID NO:31)	reverse
11	ccatcctggacttttactcc	(SEQ ID NO:32)	forward
	ctttcctgcaactgtgtttattg	(SEQ ID NO:33)	reverse

Each primer pair can be used to generate an amplified sequence of about 300 base pairs. This is especially desirable in instances in which sequence analysis is performed using SSCP gel electrophoretic procedures, in that such procedures work optimally using sequences of about 300 base pairs or less.

Additional *HKNG1* nucleic acid sequences which are preferred for such amplification-related analyses are those which will detect the presence of an *HKNG1* polymorphism which differs from the consensus *HKNG1* sequence depicted in FIG. 3A-3R. Such polymorphisms include ones which represent mutations associated with an *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia. For example, a single base mutation identified in the Example presented in Section 8, below, results in a mutant *HKNG1* gene product comprising substitution of a lysine residue for the wild-type glutamic acid residue at amino acid position 202 of the *HKNG1* amino acid sequence shown in FIG. 1A-1B (SEQ ID NO:2) or amino acid position 184 of the *HKNG1* amino acid sequence shown in FIG. 2A-2B (SEQ ID NO:4).

Additionally, well-known genotyping techniques can be performed to identify individuals carrying *HKNG1* gene mutations. Such techniques include, for example, the use of restriction fragment length polymorphisms (RFLPs), which involve sequence variations in one of the recognition sites for the specific restriction enzyme used.

Further, improved methods for analyzing DNA polymorphisms, which can be utilized for the identification of *HKNG1* gene-specific mutations, have been described that capitalize on the presence of variable numbers of short, tandemly repeated DNA sequences between the restriction enzyme sites. For example, Weber (U.S. Pat. No. 5,075,217) describes a DNA marker based on length polymorphisms in blocks of (dC-dA)n-(dG-dT)n short tandem repeats. The average separation of (dC-dA)n-(dG-dT)n blocks is estimated to be 30,000-60,000 bp. Markers that are so closely spaced exhibit a high frequency co-inheritance, and are extremely useful in the identification of genetic mutations, such as, for example, mutations within the *HKNG1* gene, and the diagnosis of diseases and disorders related to *HKNG1* mutations.

Also, Caskey *et al.* (U.S. Pat.No. 5,364,759) describe a DNA profiling assay for detecting short tri and tetra nucleotide repeat sequences. The process includes extracting

the DNA of interest, such as the *HKNG1* gene, amplifying the extracted DNA, and labelling the repeat sequences to form a genotypic map of the individual's DNA.

The level of *HKNG1* gene expression can also be assayed. For example, RNA from a cell type or tissue known, or suspected, to express the *HKNG1* gene, such as brain, may be isolated and tested utilizing hybridization or PCR techniques such as are described, above. The isolated cells can be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells to be used as part of a cell-based gene therapy technique or, alternatively, to test the effect of compounds on the expression of the *HKNG1* gene. Such analyses may reveal both quantitative and qualitative aspects of the expression pattern of the *HKNG1* gene, including activation or inactivation of *HKNG1* gene expression.

In one embodiment of such a detection scheme, a cDNA molecule is synthesized from an RNA molecule of interest (*e.g.*, by reverse transcription of the RNA molecule into cDNA). A sequence within the cDNA is then used as the template for a nucleic acid amplification reaction, such as a PCR amplification reaction, or the like. The nucleic acid reagents used as synthesis initiation reagents (*e.g.*, primers) in the reverse transcription and nucleic acid amplification steps of this method are chosen from among the *HKNG1* gene nucleic acid reagents described in Section 5.1. The preferred lengths of such nucleic acid reagents are at least 9-30 nucleotides. For detection of the amplified product, the nucleic acid amplification may be performed using radioactively or non-radioactively labeled nucleotides. Alternatively, enough amplified product may be made such that the product may be visualized by standard ethidium bromide staining or by utilizing any other suitable nucleic acid staining method.

Additionally, it is possible to perform such *HKNG1* gene expression assays "in situ", *i.e.*, directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents such as those described in Section 5.1 may be used as probes and/or primers for such in situ procedures (see, for example, Nuovo, G.J., 1992, "PCR In Situ Hybridization: Protocols And Applications", Raven Press, NY).

Alternatively, if a sufficient quantity of the appropriate cells can be obtained, standard Northern analysis can be performed to determine the level of mRNA expression of the *HKNG1* gene.

5.6.1. CHROMOSOME MAPPING

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. Accordingly, nucleic acid molecules described herein or fragments thereof, can be used to map the location of the corresponding genes on a chromosome. For example, the nucleic acid molecules described herein can be used to map the chromosomal location of *HKNG* homologues in various species. Such mapping information can be used, for example, for analysis of the activity of *HKNG* transgenes in mice. The nucleic acid molecules can further be used to map the location of copies of *HKNG* genes in the human chromosome, such as those caused by genetic abnormalties, e.g., translocations.

Briefly, genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the sequence of a gene of the invention. Computer analysis of the sequence of a gene of the invention can be used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the gene sequences will yield an amplified fragment. For a review of this technique, see D'Eustachio *et al.* (1983, Science 220:919-924).

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the nucleic acid sequences of the invention to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a gene to its chromosome include *in situ* hybridization (described in Fan et al., 1990, Proc. Natl. Acad. Sci. USA 87:6223-27), pre-screening with labeled flow-sorted chromosomes (CITE), and pre-selection by hybridization to chromosome specific cDNA libraries. Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal

spread can further be used to provide a precise chromosomal location in one step. (For a review of this technique, see Verma et al., Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York, 1988.)

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

In another embodiment, a *HKNG* polypeptide and fragments and sequences thereof and antibodies specific thereto can be used to map the location of the gene encoding the polypeptide on a chromosome. This mapping can be carried out by specifically detecting the presence of the polypeptide in members of a panel of somatic cell hybrids between cells of a first species of animal from which the protein originates and cells from a second species of animal and then determining which somatic cell hybrid(s) expresses the polypeptide and noting the chromosome(s) from the first species of animal that it contains. For examples of this technique, see Pajunen *et al.* (1988*) Cytogenet. Cell Genet. 47*:37-41 and Van Keuren *et al.* (1986*) Hum. Genet. 74*:34-40. Alternatively, the presence of the polypeptide in the somatic cell hybrids can be determined by assaying an activity or property of the polypeptide, for example, enzymatic activity, as described in Bordelon-Riser *et al.* (1979*) Somatic Cell Genetics 5*:597-613 and Owerbach *et al.* (1978*) Proc. Natl. Acad. Sci. USA 75*:5640-5644.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, *e.g.*, Egeland et al.,1987, Nature 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with a gene of the invention can be determined. If

a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

5.6.2. TISSUE TYPING

The nucleic acid sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the nucleic acid sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The nucleic acid sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500

bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of FIG. 3A-3R can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those shown in FIG. 3A-3R are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from the nucleic acid sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

5.6.3. USE OF GENE SEQUENCES IN FORENSIC BIOLOGY

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, *e.g.*, hair or skin, or body fluids, e.g., blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e. another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions are particularly appropriate for this use

as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the nucleic acid sequences of the invention or portions thereof, e.g., fragments derived from noncoding regions having a length of at least 20 or 30 bases.

The *HKNG* nucleic acid sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, e.g., brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such probes can be used to identify tissue by species and/or by organ type.

5.6.4. USE OF *HKNG* GENE SEQUENCES IN PREDICTIVE MEDICINE

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining HKNG protein and/or nucleic acid expression as well as HKNG activity, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant or unwanted HKNG expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with an HKNG protein, nucleic acid expression or activity. For example, mutations in an HKNG gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with an HKNG protein, nucleic acid expression or activity.

As an alternative to making determinations based on the absolute expression level of selected genes, determinations may be based on the normalized expression levels of these genes. Expression levels are normalized by correcting the absolute expression level of an *HKNG* gene by comparing its expression to the expression of a gene that is not an *HKNG*

gene, e.g., a housekeeping gene that is constitutively expressed. Suitable genes for normalization include housekeeping genes such as the actin gene. This normalization allows the comparison of the expression level in one sample, e.g., a patient sample, to another sample, e.g., a non-BAD-affected normal sample, or between samples from different sources.

Alternatively, the expression level can be provided as a relative expression level. To determine a relative expression level of a gene, the level of expression of the gene is determined for 10 or more samples of different cell isolates, preferably 50 or more samples, prior to the determination of the expression level for the sample in question. The cell isolates are selected depending upon the tissues in which the gene of interest is expressed. For example, for HKNG family members, expression was observed in the brain. The mean expression level of each of the genes assayed in the larger number of samples is determined and this is used as a baseline expression level for the gene(s) in question. The expression level of the gene determined for the test sample (absolute level of expression) is then divided by the mean expression value obtained for that gene. This provides a relative expression level and aids in identifying extreme cases of a HKNG-mediated disease.

For example, by way of illustration only, for HKNG family members, diseases which may be studied include, without limitation, those associated with tissues of the brain.

Preferably, the samples used in the baseline determination will be from an HKNG-mediated diseased or from non-diseased cells of tissue. The choice of the cell source is dependent on the use of the relative expression level. Using expression found in normal tissues as a mean expression score aids in validating whether the HKNG gene assayed is cell-type specific for the tissues in which expression is observed versus the expression found in normal cells. Such a use is particularly important in identifying whether an *HKNG* gene can serve as a target gene. In addition, as more data is accumulated, the mean expression value can be revised, providing improved relative expression values based on accumulated data. Expression data from brain cells provides a means for grading the severity of the HKNG-mediated disease state.

Another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of HKNG in clinical trials.

5.7. DETECTION OF *HKNG1* GENE PRODUCTS

Antibodies directed against unimpaired or mutant *HKNG1* gene products or conserved variants or peptide fragments thereof, which are discussed, above, in Section 5.3, may also be used as diagnostics and prognostics for a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia. Such methods may be used to detect abnormalities in the level of *HKNG1* gene product synthesis or expression, or abnormalities in the structure, temporal expression, and/or physical location of *HKNG1* gene product. The antibodies and immunoassay methods described herein have, for example, important *in vitro* applications in assessing the efficacy of treatments for *HKNG1*-mediated neuropsychiatric disorders, such as BAD and schizophrenia. Antibodies, or fragments of antibodies, such as those described below, may be used to screen potentially therapeutic compounds *in vitro* to determine their effects on *HKNG1* gene expression and *HKNG1* gene product production. The compounds that have beneficial effects on a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia.

In vitro immunoassays may also be used, for example, to assess the efficacy of cell-based gene therapy for a *HKNG1*-mediated neuropsychiatric disorder, such as BAD schizophrenia. Antibodies directed against *HKNG1* gene products may be used *in vitro* to determine, for example, the level of *HKNG1* gene expression achieved in cells genetically engineered to produce *HKNG1* gene product. In the case of intracellular *HKNG1* gene products, such an assessment is done, preferably, using cell lysates or extracts. Such analysis will allow for a determination of the number of transformed cells necessary to achieve therapeutic efficacy *in vivo*, as well as optimization of the gene replacement protocol.

The tissue or cell type to be analyzed will generally include those that are known, or suspected, to express the *HKNG1* gene. The protein isolation methods employed herein may, for example, be such as those described in Harlow and Lane (1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). The isolated cells can be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells to

be used as part of a cell-based gene therapy technique or, alternatively, to test the effect of compounds on the expression of the *HKNG1* gene.

Preferred diagnostic methods for the detection of *HKNG1* gene products, conserved variants or peptide fragments thereof, may involve, for example, immunoassays wherein the *HKNG1* gene products or conserved variants or peptide fragments are detected by their interaction with an anti-*HKNG1* gene product-specific antibody.

For example, antibodies, or fragments of antibodies, such as those described, above, in Section 5.3, may be used to quantitatively or qualitatively detect the presence of *HKNG1* gene products or conserved variants or peptide fragments thereof. This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody (see below, this Section) coupled with light microscopic, flow cytometric, or fluorimetric detection. Such techniques are especially preferred for *HKNG1* gene products that are expressed on the cell surface.

The antibodies (or fragments thereof) useful in the present invention may, additionally, be employed histologically, as in immunofluorescence or immunoelectron microscopy, for *in situ* detection of *HKNG1* gene products, conserved variants or peptide fragments thereof. *In situ* detection may be accomplished by removing a histological specimen from a patient, and applying thereto a labeled antibody that binds to an rTs polypeptide. The antibody (or fragment) is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the presence of the *HKNG1* gene product, conserved variants or peptide fragments, but also its distribution in the examined tissue. Using the present invention, those of ordinary skill will readily recognize that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve *in situ* detection of a *HKNG1* gene product.

Immunoassays for *HKNG1* gene products, conserved variants, or peptide fragments thereof will typically comprise incubating a sample, such as a biological fluid, a tissue extract, freshly harvested cells, or lysates of cells in the presence of a detectably labeled antibody capable of identifying *HKNG1* gene product, conserved variants or peptide fragments thereof, and detecting the bound antibody by any of a number of techniques well-known in the art. As used herein, the term "patient sample", "sample", "biological sample"

or "appropriate cellular source" refers to a sample of tissue or fluid suspected of containing a mutated or non-mutated *HKNG1* polynucleotide or polypeptide from an individual including, but not limited to, e.g., blood, plasma, serum, ascites, pleural effusion, thoracentisis, spinal fluid, lymph fluid, bone marrow, the external sections of the skin, respiratory, intestinal, and genito-urinary tracts, stool, urine, sputum, tears, saliva, blood cells, tumors, organs, tissue and samples of in vitro cell culture constituents.

Thus, in the assessment of whether a normal individual or a BAD-affected patient is expressing the secreted *HKNG1* gene product, or conserved variants or peptide fragments thereof, most simply, blood can be drawn and the blood sample incubated in the presence of a detectably labeled antibody capable of identifying *HKNG1* gene products or conserved variants or peptide fragments thereof, and detecting the bound antibody by any of a number of techniques well-known in the art. The range of the *HKNG1* gene products, or conserved variants or peptide fragments thereof, which may be detected in the blood or any one of the patient samples listed *supra* using a detectably labeled antibody capable of identifying *HKNG1* gene products or conserved variants or peptide fragments thereof is from about 1 ng/ml to about 100 ng/ml. More preferred ranges for detection of *HKNG1* gene products or conserved variants or peptide fragments thereof are about 10 ng/ml to about 90 ng/ml, about 20 ng/ml to about 80 ng/ml, about 25 ng/ml to about 70 ng/ml, and about 30 ng/ml to about 60 ng/ml. The most preferred range for the detection of *HKNG1* gene products or conserved variants or peptide fragments thereof is about 35 ng/ml to about 40 ng/ml.

The biological sample may be brought in contact with and immobilized onto a solid phase support or carrier, such as nitrocellulose, that is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the detectably labeled *HKNG1* gene product specific antibody. The solid phase support may then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on the solid support may then be detected by conventional means.

By "solid phase support or carrier" is intended any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses,

polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine experimentation.

One of the ways in which the *HKNG1* gene product-specific antibody can be detectably labeled is by linking the same to an enzyme, such as for use in an enzyme immunoassay (EIA) (Voller, A., "The Enzyme Linked Immunosorbent Assay (ELISA)", 1978, Diagnostic Horizons 2:1-7, Microbiological Associates Quarterly Publication, Walkersville, MD); Voller, A. *et al.*, 1978, J. Clin. Pathol. 31:507-520; Butler, J.E., 1981, Meth. Enzymol. 73:482-523; Maggio, E. (ed.), 1980, Enzyme Immunoassay, CRC Press, Boca Raton, FL,; Ishikawa, E. *et al.*, (eds.), 1981, Enzyme Immunoassay, Kgaku Shoin, Tokyo). The enzyme which is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety that can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes that can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, α-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, β-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods that employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect *HKNG1* gene products through the use of a

radioimmunoassay (RIA) (see, for example, Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, *o*-phthaldehyde and fluorescamine.

The antibody can also be detectably labeled using fluorescence emitting metals such as ^{152}Eu, or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

Further, an antibody (or fragment thereof) can be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone,

glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, α-interferon, β-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophase colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982).

Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

Accordingly, in one aspect, the invention provides substantially purified antibodies or fragment thereof, and non-human antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of any one of SEQ ID Nos. 2 or 4; a fragment of at least 15 amino acid residues of the amino acid sequence of any one of SEQ ID Nos. 2 or 4, an amino acid sequence which is at least 95% identical to the amino acid sequence of any one of SEQ ID Nos. 2 or 4, wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of any one of SEQ ID Nos. 1, 3, 5, 6, or 7, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. In various embodiments, the substantially purified antibodies of the invention, or fragments thereof, can be human, non-human, chimeric and/or humanized antibodies.

In another aspect, the invention provides non-human antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of any one of SEQ ID Nos. 2 or 4; a fragment of at least 15 amino acid residues of the amino acid sequence of any one of SEQ ID Nos. 2 or 4, an amino acid sequence which is at least 95% identical to the amino acid sequence of any one of SEQ ID Nos. 2 or 4, wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of any one of SEQ ID Nos. 1, 3, 5, 6, or 7, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. Such non-human antibodies can be goat, mouse, sheep, horse, chicken, rabbit, or rat antibodies. Alternatively, the non-human antibodies of the invention can be chimeric and/or humanized antibodies. In addition, the non-human antibodies of the invention can be polyclonal antibodies or monoclonal antibodies.

In still a further aspect, the invention provides monoclonal antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of any one of SEQ ID Nos. 2 or 4; a fragment of at least 15 amino acid residues of the amino acid sequence of any one of SEQ ID Nos. 2 or 4, an amino acid sequence which is at least 95% identical to the amino acid sequence of any one of SEQ ID Nos. 2 or 4, wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of any one of SEQ ID Nos. 1, 3, 5, 6, or 7, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. The monoclonal antibodies can be human, humanized, chimeric and/or non-human antibodies.

The substantially purified antibodies or fragments thereof specifically bind to a signal peptide, a secreted sequence, an extracellular domain, a transmembrane or a cytoplasmic domain cytoplasmic membrane of a polypeptide of the invention. In one embodiment, the substantially purified antibodies or fragments thereof, the human or non-human antibodies or fragments thereof, and/or the monoclonal antibodies or fragments thereof, of the invention specifically bind to a secreted sequence or an extracellular domain of the amino acid sequence of SEQ ID Nos. 2 or 4.

Any of the antibodies of the invention can be conjugated to a therapeutic moiety or to a detectable substance. Non-limiting examples of detectable substances that can be conjugated to the antibodies of the invention are an enzyme, a prosthetic group, a fluorescent material, a luminescent material, a bioluminescent material, and a radioactive material.

The invention also provides a kit containing an antibody of the invention conjugated to a detectable substance, and instructions for use. Still another aspect of the invention is a pharmaceutical composition comprising an antibody of the invention and a pharmaceutically acceptable carrier. In one embodiment, the pharmaceutical composition contains an antibody of the invention, a therapeutic moiety, and a pharmaceutically acceptable carrier.

Still another aspect of the invention is a method of making an antibody that specifically recognizes HKNG, the method comprising immunizing a mammal with a polypeptide. The polypeptide used as an immungen comprises an amino acid sequence selected from the group consisting of: the amino acid sequence of any one of SEQ ID Nos. 2 or 4; a fragment of at least 15 amino acid residues of the amino acid sequence of any one of SEQ ID Nos. 2 or 4, an amino acid sequence which is at least 95% identical to the amino acid sequence of any one of SEQ ID Nos. 2 or 4, wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of any one of SEQ ID Nos. 1, 3, 5, 6, 7, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. After immunization, a sample is collected from the mammal that contains an antibody that specifically recognizes a HKNG polypeptide as exemplified in SEQ ID Nos. 2 or 4, or portions thereof. Preferably, the polypeptide is recombinantly produced using a non-human host cell. Optionally, the antibodies can be further purified from the sample using techniques well known to those of skill in the art. The method can further comprise producing a monoclonal antibody-producing cell from the cells of the mammal. Optionally, antibodies are collected from the antibody-producing cell.

5.8. SCREENING ASSAYS FOR COMPOUNDS THAT MODULATE *HKNG1* GENE ACTIVITY

The following assays are designed to identify compounds that bind to a *HKNG1* gene product, compounds that bind to proteins, or portions of proteins that interact with a *HKNG1* gene product, compounds that interfere with the interaction of a *HKNG1* gene product with proteins and compounds that modulate the activity of the *HKNG1* gene (*i.e.*, modulate the level of *HKNG1* gene expression and/or modulate the level of *HKNG1* gene product activity). Assays may additionally be utilized that identify compounds that bind to *HKNG1* gene regulatory sequences (*e.g.*, promoter sequences; see *e.g.*, Platt, 1994, J. Biol. Chem. 269, 28558-28562), and that can modulate the level of *HKNG1* gene expression. Such compounds may include, but are not limited to, small organic molecules,

such as ones that are able to cross the blood-brain barrier, gain to and/or entry into an appropriate cell and affect expression of the *HKNG1* gene or some other gene involved in a *HKNG1* regulatory pathway, such as *LRP-2* or *RAP*, or intracellular proteins.

Methods for the identification of such proteins are described, below, in Section 5.8.2. Such proteins may be involved in the control and/or regulation of mood. Further, among these compounds are compounds that affect the level of *HKNG1* gene expression and/or *HKNG1* gene product activity and that can be used in the therapeutic treatment of *HKNG1*-mediated neuropsychiatric disorders such as BAD and schizophrenia as described, below, in Section 5.9.

Compounds may include, but are not limited to, peptides such as, for example, soluble peptides, including but not limited to, Ig-tailed fusion peptides, and members of random peptide libraries; (see, *e.g.*, Lam, *et al.*, 1991, Nature 354:82-84; Houghten, *et al.*, 1991, Nature 354:84-86), and combinatorial chemistry-derived molecular library made of D- and/or L- configuration amino acids, phosphopeptides (including, but not limited to members of random or partially degenerate, directed phosphopeptide libraries; see, *e.g.*, Songyang, *et al.*, 1993, Cell 72:767-778), antibodies (including, but not limited to, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and FAb, $F(ab')_2$ and FAb expression library fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules.

Such compounds may further comprise compounds, in particular drugs or members of classes or families of drugs, known to ameliorate or exacerbate the symptoms of a neuropsychiatric disorder such as BAD or schizophrenia. Such compounds include antidepressants such as lithium salts, carbamazepine, valproic acid, lysergic acid diethylamide (LSD), *p*-chlorophenylalanine, *p*-propyldopacetamide dithiocarbamate derivatives *e.g.*, FLA 63; anti-anxiety drugs, *e.g.*, diazepam; monoamine oxidase (MAO) inhibitors, *e.g.*, iproniazid, clorgyline, phenelzine and isocarboxazid; biogenic amine uptake blockers, *e.g.*, tricyclic antidepressants such as desipramine, imipramine and amitriptyline; serotonin reuptake inhibitors *e.g.*, fluoxetine; antipsychotic drugs such as phenothiazine derivatives (*e.g.*, chlorpromazine (thorazine) and trifluopromazine)), butyrophenones (*e.g.*, haloperidol (Haldol)), thioxanthene derivatives (*e.g.*, chlorprothixene), and dibenzodiazepines (*e.g.*, clozapine); benzodiazepines; dopaminergic agonists and

antagonists *e.g.*, L-DOPA, cocaine, amphetamine, α-methyl-tyrosine, reserpine, tetrabenazine, benzotropine, pargyline; noradrenergic agonists and antagonists *e.g.*, clonidine, phenoxybenzamine, phentolamine, tropolone.

Compounds identified via assays such as those described herein may be useful, for example, in elaborating the biological function of the *HKNG1* gene product and for ameliorating *HKNG1*-mediated neuropsychiatric disorders, such as BAD and schizophrenia. Assays for testing the effectiveness of compounds identified by, for example, techniques such as those described in Sections 5.8.1 - 5.8.3, are discussed, below, in Section 5.8.4.

5.8.1. *IN VITRO* SCREENING ASSAYS FOR COMPOUNDS THAT BIND TO THE *HKNG1* GENE PRODUCT

In vitro systems may be designed to identify compounds capable of binding the *HKNG1* gene products of the invention. Compounds identified may be useful, for example, in modulating the activity of unimpaired and/or mutant *HKNG1* gene products, may be useful in elaborating the biological function of the *HKNG1* gene product, may be utilized in screens for identifying compounds that disrupt normal *HKNG1* gene product interactions, or may in themselves disrupt such interactions.

The principle of the assays used to identify compounds that bind to the *HKNG1* gene product involves preparing a reaction mixture of the *HKNG1* gene product and the test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus forming a complex that can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways. For example, one method to conduct such an assay involves anchoring a *HKNG1* gene product or a test substance onto a solid support and detecting *HKNG1* gene product/test compound complexes formed on the solid support at the end of the reaction. In one embodiment of such a method, the *HKNG1* gene product may be anchored onto a solid support, and the test compound, which is not anchored, may be labeled, either directly or indirectly.

In practice, microtiter plates are conveniently utilized as the solid support. The anchored component may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished by simply coating the solid surface with a

solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal antibody, specific for the protein to be immobilized may be used to anchor the protein to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the non-immobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the previously non-immobilized component (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody).

Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; *e.g.*, using an immobilized antibody specific for *HKNG1* gene product or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

5.8.2. ASSAYS FOR INTRACELLULAR PROTEINS THAT INTERACT WITH *HKNG1* GENE PRODUCTS

Any method suitable for detecting protein-protein interactions may be employed for identifying *HKNG1* gene product-protein interactions.

Among the traditional methods that may be employed are co-immunoprecipitation, cross-linking and co-purification through gradients or chromatographic columns. Utilizing procedures such as these allows for the identification of proteins, including intracellular proteins, that interact with *HKNG1* gene products. Once isolated, such a protein can be identified and can be used in conjunction with standard techniques, to identify proteins it interacts with. For example, at least a portion of the amino acid sequence of a protein that interacts with the *HKNG1* gene product can be

ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique (see, *e.g.*, Creighton, 1983, "Proteins: Structures and Molecular Principles," W.H. Freeman & Co., N.Y., pp.34-49). The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for gene sequences encoding such proteins. Screening made be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and the screening are well-known. (See, *e.g.*, Ausubel, *supra*, and 1990, "PCR Protocols: A Guide to Methods and Applications," Innis, *et al.*, eds. Academic Press, Inc., New York).

Additionally, methods may be employed that result in the simultaneous identification of genes that encode a protein which interacts with a *HKNG1* gene product. These methods include, for example, probing expression libraries with labeled *HKNG1* gene product, using *HKNG1* gene product in a manner similar to the well known technique of antibody probing of λgt11 libraries.

One method that detects protein interactions *in vivo*, the two-hybrid system, is described in detail for illustration only and not by way of limitation. One version of this system has been described (Chien, *et al.*, 1991, Proc. Natl. Acad. Sci. USA, 88:9578-9582) and is commercially available from Clontech (Palo Alto, CA).

Briefly, utilizing such a system, plasmids are constructed that encode two hybrid proteins: one consists of the DNA-binding domain of a transcription activator protein fused to the *HKNG1* gene product and the other consists of the transcription activator protein's activation domain fused to an unknown protein that is encoded by a cDNA that has been recombined into this plasmid as part of a cDNA library. The DNA-binding domain fusion plasmid and the cDNA library are transformed into a strain of the yeast *Saccharomyces cerevisiae* that contains a reporter gene (*e.g.*, HBS or *lacZ*) whose regulatory region contains the transcription activator's binding site. Either hybrid protein alone cannot activate transcription of the reporter gene: the DNA-binding domain hybrid cannot because it does not provide activation function and the activation domain hybrid cannot because it cannot localize to the activator's binding sites. Interaction of the two hybrid proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The two-hybrid system or related methodologies may be used to screen activation domain libraries for proteins that interact with the "bait" gene product. By way of example, and not by way of limitation, *HKNG1* gene products may be used as the bait gene product. Total genomic or cDNA sequences are fused to the DNA encoding an activation domain. This library and a plasmid encoding a hybrid of a bait *HKNG1* gene product fused to the DNA-binding domain are co-transformed into a yeast reporter strain, and the resulting transformants are screened for those that express the reporter gene. For example, a bait *HKNG1* gene sequence, such as the open reading frame of the *HKNG1* gene, can be cloned into a vector such that it is translationally fused to the DNA encoding the DNA-binding domain of the GAL4 protein. These colonies are purified and the library plasmids responsible for reporter gene expression are isolated. DNA sequencing is then used to identify the proteins encoded by the library plasmids.

A cDNA library of the cell line from which proteins that interact with bait *HKNG1* gene product are to be detected can be made using methods routinely practiced in the art. According to the particular system described herein, for example, the cDNA fragments can be inserted into a vector such that they are translationally fused to the transcriptional activation domain of GAL4. Such a library can be co-transformed along with the bait *HKNG1* gene-GAL4 fusion plasmid into a yeast strain that contains a lacZ gene driven by a promoter that contains GAL4 activation sequence. A cDNA encoded protein, fused to a GAL4 transcriptional activation domain that interacts with bait *HKNG1* gene product will reconstitute an active GAL4 protein and thereby drive expression of the HIS3 gene. Colonies that express HIS3 can be detected by their growth on petri dishes containing semi-solid agar based media lacking histidine. The cDNA can then be purified from these strains, and used to produce and isolate the bait *HKNG1* gene product-interacting protein using techniques routinely practiced in the art.

5.8.3. ASSAYS FOR COMPOUNDS THAT INTERFERE WITH *HKNG1* GENE PRODUCT MACROMOLECULE INTERACTION

The *HKNG1* gene products may, *in vivo*, interact with one or more macromolecules, including intracellular macromolecules, such as proteins. Such macromolecules may include, but are not limited to, the receptor LRP-2 (also known as megalin or gp330) and the receptor-associated protein (RAP) which are described in Section 9 below. Additional macromolecules which interact with *HKNG1* gene products *in vivo* may include nucleic acid molecules and those proteins identified via methods such as those described, above, in Sections 5.8.1 - 5.8.2. For purposes of this discussion, the macromolecules are referred to herein as "binding partners". Compounds such as RAP that disrupt *HKNG1* gene product binding to a binding partner may be useful in regulating the activity of the *HKNG1* gene product, especially mutant *HKNG1* gene products. Besides RAP, such compounds may also include, but are not limited to molecules such as peptides, and the like, as described, for example, in Section 5.8.2 above.

The basic principle of an assay system used to identify compounds that interfere with the interaction between the *HKNG1* gene product and a binding partner or partners involves preparing a reaction mixture containing the *HKNG1* gene product and the binding partner (for example LRP-2) under conditions and for a time sufficient to allow the two to interact and bind, thus forming a complex. In order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound may be initially included in the reaction mixture, or may be added at a time subsequent to the addition of *HKNG1* gene product and its binding partner. Control reaction mixtures are incubated without the test compound or with a compound which is known not to block complex formation. The formation of any complexes between the *HKNG1* gene product and the binding partner is then detected. The formation of a complex in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the *HKNG1* gene product and the binding partner. Additionally, complex formation within reaction mixtures containing the test compound and normal *HKNG1* gene product may also be compared to complex formation within reaction mixtures containing the test compound and a mutant *HKNG1* gene product. This comparison may be important in those cases wherein it is desirable to

identify compounds that disrupt interactions of mutant but not normal *HKNG1* gene product.

In an alternative embodiment, the above assay may be performed using a reaction mixture containing the *HKNG1* gene product, a binding partner, such as LRP-2, and a third compound, such as RAP, which disrupts *HKNG1* gene product binding to the binding partner. The reaction mixture is prepared and incubated in the presence and absence of the test compound, as described above, and the formation of any complexes between the *HKNG1* gene product and the binding partner is detected. In this embodiment, the formation of a complex in the reaction mixture containing the test compound, but not in the control reaction, indicates that the test compound interferes with the ability of the second compound to disrupt *HKNG1* gene product binding to its binding partner.

The assay for compounds that interfere with the interaction of the *HKNG1* gene products and binding partners can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the *HKNG1* gene product or the binding partner onto a solid support and detecting complexes formed on the solid support at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the *HKNG1* gene products and the binding partners, *e.g.*, by competition, can be identified by conducting the reaction in the presence of the test substance; *i.e.*, by adding the test substance to the reaction mixture prior to or simultaneously with the *HKNG1* gene product and interactive intracellular binding partner. Alternatively, test compounds that disrupt preformed complexes, *e.g.*, compounds with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are described briefly below.

In a heterogeneous assay system, either the *HKNG1* gene product or the interactive binding partner, is anchored onto a solid surface, while the non-anchored species is labeled, either directly or indirectly. In practice, microtiter plates are conveniently utilized. The anchored species may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid

surface with a solution of the *HKNG1* gene product or binding partner and drying. Alternatively, an immobilized antibody specific for the species to be anchored may be used to anchor the species to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds that inhibit complex formation or that disrupt preformed complexes can be detected.

Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; *e.g.*, using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds that inhibit complex formation or that disrupt preformed complexes can be identified.

In an alternate embodiment of the invention, a homogeneous assay can be used. In this approach, a preformed complex of the *HKNG1* gene product and the interactive binding partner is prepared in which either the *HKNG1* gene product or its binding partners is labeled, but the signal generated by the label is quenched due to complex formation (see, *e.g.*, U.S. Patent No. 4,109,496 by Rubenstein which utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above

background. In this way, test substances that disrupt *HKNG1* gene product/binding partner interaction can be identified.

In another embodiment of the invention, these same techniques can be employed using peptide fragments that correspond to the binding domains of the *HKNG1* product and/or the binding partner (in cases where the binding partner is a protein), in place of one or both of the full length proteins. Any number of methods routinely practiced in the art can be used to identify and isolate the binding sites. These methods include, but are not limited to, mutagenesis of the gene encoding one of the proteins and screening for disruption of binding in a co-immunoprecipitation assay. Compensating mutations in the gene encoding the second species in the complex can then be selected. Sequence analysis of the genes encoding the respective proteins will reveal the mutations that correspond to the region of the protein involved in interactive binding. Alternatively, one protein can be anchored to a solid surface using methods described in this Section above, and allowed to interact with and bind to its labeled binding partner, which has been treated with a proteolytic enzyme, such as trypsin. After washing, a short, labeled peptide comprising the binding domain may remain associated with the solid material, which can be isolated and identified by amino acid sequencing. Also, once the gene coding for the segments is engineered to express peptide fragments of the protein, it can then be tested for binding activity and purified or synthesized.

For example, and not by way of limitation, a *HKNG1* gene product can be anchored to a solid material as described, above, in this Section by making a GST-*HKNG1* fusion protein and allowing it to bind to glutathione agarose beads. The binding partner can be labeled with a radioactive isotope, such as ^{35}S, and cleaved with a proteolytic enzyme such as trypsin. Cleavage products can then be added to the anchored GST-*HKNG1* fusion protein and allowed to bind. After washing away unbound peptides, labeled bound material, representing the binding partner binding domain, can be eluted, purified, and analyzed for amino acid sequence by well-known methods. Peptides so identified can be produced synthetically or produced using recombinant DNA technology.

5.8.4. ASSAYS FOR IDENTIFICATION OF COMPOUNDS THAT AMELIORATE A *HKNG1*-MEDIATED NEUROPSYCHIATRIC DISORDERS

Compounds, including but not limited to binding compounds identified via assay techniques such as those described, above, in Sections 5.8.1 - 5.8.4, can be tested for the ability to ameliorate symptoms of a *HKNG1*-mediated neuropsychiatric disorder, such as schizophrenia and bipolar affective (mood) disorders, including severe bipolar affective (mood) disorder (BP-I), bipolar affective (mood) disorder with hypomania and major depression (BP-II).

It should be noted that the assays described herein can identify compounds that affect *HKNG1* activity by either affecting *HKNG1* gene expression or by affecting the level of *HKNG1* gene product activity. For example, compounds may be identified that are involved in another step in the pathway in which the *HKNG1* gene and/or *HKNG1* gene product is involved and, by affecting this same pathway may modulate the effect of HKNG1 on the development of a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia. Such compounds can be used as part of a therapeutic method for the treatment of the disorder.

Described below are cell-based and animal model-based assays for the identification of compounds exhibiting such an ability to ameliorate symptoms of a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia.

First, cell-based systems can be used to identify compounds that may act to ameliorate symptoms of a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia. Such cell systems can include, for example, recombinant or non-recombinant cell, such as cell lines, that express the *HKNG1* gene.

In utilizing such cell systems, cells that express *HKNG1* may be exposed to a compound suspected of exhibiting an ability to ameliorate symptoms of a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia, at a sufficient concentration and for a sufficient time to elicit such an amelioration of such symptoms in the exposed cells. After exposure, the cells can be assayed to measure alterations in the expression of the *HKNG1* gene, *e.g.*, by assaying cell lysates for *HKNG1* mRNA transcripts (*e.g.*, by Northern analysis) or for *HKNG1* gene products expressed by the cell; compounds that modulate expression of the *HKNG1* gene are good candidates as therapeutics.

In addition, animal-based systems or models for a *HKNG1*-mediated neuropsychiatric disorder, for example, transgenic mice containing a human or altered form of *HKNG1* gene, may be used to identify compounds capable of ameliorating symptoms of the disorder. Such animal models may be used as test substrates for the identification of drugs, pharmaceuticals, therapies and interventions. For example, animal models may be exposed to a compound suspected of exhibiting an ability to ameliorate symptoms, at a sufficient concentration and for a sufficient time to elicit such an amelioration of symptoms of a *HKNG1* neuropsychiatric disorder. The response of the animals to the exposure may be monitored by assessing the reversal of the symptoms of the disorder.

With regard to intervention, any treatments that reverse any aspect of symptoms of a *HKNG1*-mediated neuropsychiatric disorder, should be considered as candidates for human therapeutic intervention in such a disorder. Dosages of test agents may be determined by deriving dose-response curves, as discussed in Section 5.10.1, below.

5.9. COMPOUNDS AND METHODS FOR THE TREATMENT OF *HKNG1*-MEDIATED NEUROPSYCHIATRIC DISORDERS

Described below are methods and compositions whereby a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia, may be treated. For example, such methods can comprise administering compounds which modulate the expression of a mammalian *HKNG1* gene and/or the synthesis or activity of a mammalian *HKNG1* gene product so symptoms of the disorder are ameliorated.

Alternatively, in those instances whereby the mammalian *HKNG1*-mediated neuropsychiatric disorders result from *HKNG1* gene mutations, such methods can comprise supplying the mammal with a nucleic acid molecule encoding an unimpaired *HKNG1* gene product such that an unimpaired *HKNG1* gene product is expressed and symptoms of the disorder are ameliorated.

In another embodiment of methods for the treatment of mammalian *HKNG1*-mediated neuropsychiatric disorders resulting from *HKNG1* gene mutations, such methods can comprise supplying the mammal with a cell comprising a nucleic acid molecule that encodes an unimpaired *HKNG1* gene product such that the cell expresses the unimpaired *HKNG1* gene product and symptoms of the disorder are ameliorated.

In cases in which a loss of normal *HKNG1* gene product function results in the development of a *HKNG1*-mediated neuropsychiatric disorder an increase in *HKNG1* gene product activity would facilitate progress towards an asymptomatic state in individuals exhibiting a deficient level of *HKNG1* gene expression and/or *HKNG1* gene product activity. Methods for enhancing the expression or synthesis of *HKNG1* can include, for example, methods such as those described below, in Section 5.9.2.

Alternatively, symptoms of *HKNG1*-mediated neuropsychiatric disorders, may be ameliorated by administering a compound that decreases the level of *HKNG1* gene expression and/or *HKNG1* gene product activity. Methods for inhibiting or reducing the level of *HKNG1* gene product synthesis or expression can include, for example, methods such as those described in Section 5.9.1.

In one embodiment of treatment methods, the compounds administered comprise compounds, in particular drugs, reported to ameliorate or exacerbate the symptoms of a neuropsychiatric disorder, such as BAD or schizophrenia. Such compounds include antidepressants such as lithium salts, carbamazepine, valproic acid, lysergic acid diethylamide (LSD), *p*-chlorophenylalanine, *p*-propyldopacetamide dithiocarbamate derivatives *e.g.*, FLA 63; anti-anxiety drugs, *e.g.*, diazepam; monoamine oxidase (MAO) inhibitors, *e.g.*, iproniazid, clorgyline, phenelzine and isocarboxazid; biogenic amine uptake blockers, *e.g.*, tricyclic antidepressants such as desipramine, imipramine and amitriptyline; serotonin reuptake inhibitors *e.g.*, fluoxetine; antipsychotic drugs such as phenothiazine derivatives (*e.g.*, chlorpromazine (thorazine) and trifluopromazine)), butyrophenones (*e.g.*, haloperidol (Haldol)), thioxanthene derivatives (*e.g.*, chlorprothixene), and dibenzodiazepines (*e.g.*, clozapine); benzodiazepines; dopaminergic agonists and antagonists *e.g.*, L-DOPA, cocaine, amphetamine, α-methyl-tyrosine, reserpine, tetrabenazine, benzotropine, pargyline; noradrenergic agonists and antagonists *e.g.*, clonidine, phenoxybenzamine, phentolamine, tropolone.

5.9.1. INHIBITORY ANTISENSE, RIBOZYME AND TRIPLE HELIX APPROACHES

In another embodiment, symptoms of *HKNG1*-mediated neuropsychiatric disorders may be ameliorated by decreasing the level of *HKNG1* gene expression and/or

HKNG1 gene product activity by using *HKNG1* gene sequences in conjunction with well-known antisense, gene "knock-out," ribozyme and/or triple helix methods to decrease the level of *HKNG1* gene expression. Among the compounds that may exhibit the ability to modulate the activity, expression or synthesis of the *HKNG1* gene, including the ability to ameliorate the symptoms of a *HKNG1*-mediated neuropsychiatric disorder, such as BAD or schizophrenia, are antisense, ribozyme, and triple helix molecules. Such molecules may be designed to reduce or inhibit either unimpaired, or if appropriate, mutant target gene activity. Techniques for the production and use of such molecules are well known to those of skill in the art.

Antisense RNA and DNA molecules act to directly block the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense approaches involve the design of oligonucleotides that are complementary to a target gene mRNA. The antisense oligonucleotides will bind to the complementary target gene mRNA transcripts and prevent translation. Absolute complementarily, although preferred, is not required.

A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarily to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarily and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

In one embodiment, oligonucleotides complementary to non-coding regions of the *HKNG1* gene could be used in an antisense approach to inhibit translation of endogenous *HKNG1* mRNA. Antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides or at least 50 nucleotides.

Regardless of the choice of target sequence, it is preferred that *in vitro* studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger, *et al.*, 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre, *et al.*, 1987, Proc. Natl. Acad. Sci. U.S.A. 84:648-652; PCT Publication No. WO88/09810, published December 15, 1988) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO89/10134, published April 25, 1988), hybridization-triggered cleavage agents (see, *e.g.*, Krol *et al.*, 1988, BioTechniques 6:958-976) or intercalating agents (see, *e.g.*, Zon, 1988, Pharm. Res. 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine,

2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5′-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an α-anomeric oligonucleotide. An α-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gautier, *et al.*, 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2′-0-methylribonucleotide (Inoue, *et al.*, 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue, *et al.*, 1987, FEBS Lett. 215:327-330).

Oligonucleotides of the invention may be synthesized by standard methods known in the art, *e.g.*, by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein, *et al.* (1988, Nucl. Acids Res. 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin, *et al.*, 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc.

While antisense nucleotides complementary to the target gene coding region sequence could be used, those complementary to the transcribed, untranslated region are most preferred.

Antisense molecules should be delivered to cells that express the target gene *in vivo*. A number of methods have been developed for delivering antisense DNA or RNA to cells; *e.g.*, antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (*e.g.*, antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systemically.

However, it is often difficult to achieve intracellular concentrations of the antisense sufficient to suppress translation of endogenous mRNAs. Therefore a preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfect target cells in the patient will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous target gene transcripts and thereby prevent translation of the target gene mRNA. For example, a vector can be introduced *e.g.*, such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3′ long terminal repeat of Rous sarcoma virus (Yamamoto, *et al.*, 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner, *et al.*, 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster, *et al.*, 1982, Nature 296:39-42), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct which can be introduced directly into the tissue site. Alternatively, viral vectors can be used that selectively infect the desired tissue, in which case administration may be accomplished by another route (*e.g.*, systemically).

Ribozyme molecules designed to catalytically cleave target gene mRNA transcripts can also be used to prevent translation of target gene mRNA and, therefore, expression of target gene product. (See, *e.g.*, PCT International Publication WO90/11364, published October 4, 1990; Sarver, *et al.*, 1990, Science 247, 1222-1225).

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. (For a review, see Rossi, 1994, Current Biology 4:469-471). The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage event. The composition of ribozyme molecules must include one or more sequences complementary to the target gene mRNA, and must include the well known catalytic sequence responsible for mRNA cleavage. For this sequence, see, *e.g.*, U.S. Patent No. 5,093,246, which is incorporated herein by reference in its entirety.

While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy target gene mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Myers, 1995, *Molecular Biology and Biotechnology: A Comprehensive Desk Reference*, VCH Publishers, New York, (see especially Figure 4, page 833) and in Haseloff and Gerlach, 1988, Nature, 334:585-591, which is incorporated herein by reference in its entirety.

Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the target gene mRNA, *i.e.*, to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one that occurs naturally in *Tetrahymena thermophila* (known as the IVS, or L-19 IVS RNA) and that has been extensively described by Thomas Cech and collaborators (Zaug, *et al.*, 1984, Science, 224:574-578; Zaug and Cech, 1986, Science, 231:470-475; Zaug, *et al.*, 1986, Nature, 324:429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and

Cech, 1986, Cell, 47:207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in the target gene.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (*e.g.*, for improved stability, targeting, etc.) and should be delivered to cells that express the target gene *in vivo*. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous target gene messages and inhibit translation. Because ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Endogenous target gene expression can also be reduced by inactivating or "knocking out" the target gene or its promoter using targeted homologous recombination (*e.g.*, see Smithies, *et al.*, 1985, Nature 317:230-234; Thomas and Capecchi, 1987, Cell 51:503-512; Thompson, *et al.*, 1989, Cell 5:313-321; each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional target gene (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous target gene (either the coding regions or regulatory regions of the target gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express the target gene *in vivo*. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the target gene. Such approaches are particularly suited in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive target gene (*e.g.*, see Thomas and Capecchi, 1987 and Thompson, 1989, *supra*). However this approach can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors.

Alternatively, endogenous target gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the target gene (*i.e.*, the target gene promoter and/or enhancers) to form triple helical structures that prevent transcription of the target gene in target cells in the body. (See generally, Helene, 1991,

Anticancer Drug Des., 6(6):569-584; Helene, *et al.*, 1992, Ann. N.Y. Acad. Sci., 660:27-36; and Maher, 1992, Bioassays 14(12):807-815).

Nucleic acid molecules to be used in triplex helix formation for the inhibition of transcription should be single stranded and composed of deoxynucleotides. The base composition of these oligonucleotides must be designed to promote triple helix formation via Hoogsteen base pairing rules, which generally require sizeable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC^+ triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarily to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, contain a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in GGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

In instances wherein the antisense, ribozyme, and/or triple helix molecules described herein are utilized to inhibit mutant gene expression, it is possible that the technique may so efficiently reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles that the possibility may arise wherein the concentration of normal target gene product present may be lower than is necessary for a normal phenotype. In such cases, to ensure that substantially normal levels of target gene activity are maintained, therefore, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity may, be introduced into cells via gene therapy methods such as those described, below, in Section 5.9.2 that do not contain sequences susceptible to whatever antisense, ribozyme, or triple helix treatments are being utilized. Alternatively, in instances whereby

the target gene encodes an extracellular protein, it may be preferable to co-administer normal target gene protein in order to maintain the requisite level of target gene activity.

Anti-sense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules, as discussed above. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors that incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

5.9.2. GENE REPLACEMENT THERAPY

HKNG1 gene nucleic acid sequences, described above in Section 5.1, can be utilized for the treatment of a *HKNG1*-mediated neuropsychiatric disorder. Such treatment can be in the form of gene replacement therapy. Specifically, one or more copies of a normal *HKNG1* gene or a portion of the *HKNG1* gene that directs the production of a H*KNG1* gene product exhibiting normal *HKNG1* gene function, may be inserted into the appropriate cells within a patient, using vectors that include, but are not limited to adenovirus, adeno-associated virus, and retrovirus vectors, in addition to other particles that introduce DNA into cells, such as liposomes.

Because the *HKNG1* gene is expressed in the brain, such gene replacement therapy techniques should be capable delivering *HKNG1* gene sequences to these cell types within patients. Thus, in one embodiment, techniques that are well known to those of skill in the art (see, *e.g.*, PCT Publication No. WO89/10134, published April 25, 1988) can be used to enable *HKNG1* gene sequences to cross the blood-brain barrier readily and to deliver the sequences to cells in the brain. With respect to delivery that is capable of crossing the blood-brain barrier, viral vectors such as, for example, those described above, are preferable.

In another embodiment, techniques for delivery involve direct administration of such *HKNG1* gene sequences to the site of the cells in which the *HKNG1* gene sequences are to be expressed.

Additional methods that may be utilized to increase the overall level of *HKNG1* gene expression and/or *HKNG1* gene product activity include using targeted homologous recombination methods, discussed in Section 5.2, above, to modify the expression characteristics of an endogenous *HKNG1* gene in a cell or microorganism by inserting a heterologous DNA regulatory element such that the inserted regulatory element is operatively linked with the endogenous *HKNG1* gene in question. Targeted homologous recombination can thus be used to activate transcription of an endogenous *HKNG1* gene that is "transcriptionally silent", *i.e.*, is not normally expressed or is normally expressed at very low levels, or to enhance the expression of an endogenous *HKNG1* gene that is normally expressed.

Further, the overall level of *HKNG1* gene expression and/or *HKNG1* gene product activity may be increased by the introduction of appropriate *HKNG1*-expressing cells, preferably autologous cells, into a patient at positions and in numbers that are sufficient to ameliorate the symptoms of a *HKNG1*-mediated neuropsychiatric disorder. Such cells may be either recombinant or non-recombinant.

Among the cells that can be administered to increase the overall level of *HKNG1* gene expression in a patient are normal cells, preferably brain cells, that express the *HKNG1* gene. Alternatively, cells, preferably autologous cells, can be engineered to express *HKNG1* gene sequences, and may then be introduced into a patient in positions appropriate for the amelioration of the symptoms of a *HKNG1*-mediated neuropsychiatric disorder. Alternately, cells that express an unimpaired *HKNG1* gene and that are from a MHC matched individual can be utilized, and may include, for example, brain cells. The expression of the *HKNG1* gene sequences is controlled by the appropriate gene regulatory sequences to allow such expression in the necessary cell types. Such gene regulatory sequences are well known to the skilled artisan. Such cell-based gene therapy techniques are well known to those skilled in the art, see, *e.g.*, Anderson, U.S. Patent No. 5,399,349.

When the cells to be administered are non-autologous cells, they can be administered using well known techniques that prevent a host immune response against the

introduced cells from developing. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Additionally, compounds, such as those identified via techniques such as those described, above, in Section 5.8, that are capable of modulating *HKNG1* gene product activity can be administered using standard techniques that are well known to those of skill in the art. In instances in which the compounds to be administered are to involve an interaction with brain cells, the administration techniques should include well known ones that allow for a crossing of the blood-brain barrier.

5.10. PHARMACEUTICAL PREPARATIONS AND METHODS OF ADMINISTRATION

The compounds that are determined to affect *HKNG1* gene expression or gene product activity can be administered to a patient at therapeutically effective doses to treat or ameliorate a *HKNG1*-mediated neuropsychiatric disorder. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of such a disorder.

5.10.1. EFFECTIVE DOSE

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds that exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies

preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (*i.e.*, the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

5.10.2. FORMULATIONS AND USE

Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients.

Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (*e.g.*, pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (*e.g.*, lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (*e.g.*, magnesium stearate, talc or silica); disintegrants (*e.g.*, potato starch or sodium starch glycolate); or wetting agents (*e.g.*, sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (*e.g.*, sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (*e.g.*, lecithin or acacia); non-aqueous vehicles (*e.g.*, almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (*e.g.*, methyl or

propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

The compositions may, if desired, be presented in a pack or dispenser device that may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

6. EXAMPLE: THE *HKNG1* GENE OF CHROMOSOME 18 IS ASSOCIATED WITH THE NEUROPSYCHIATRIC DISORDER BAD

In the Example presented in this Section, studies are described that define a narrow interval of approximately 25 kb on the short arm of human chromosome 18 which is associated with the neuropsychiatric disorder BAD. The interval is demonstrated to lie within the gene referred to herein as the *HKNG1* gene.

6.1. MATERIALS AND METHODS

6.1.1. LINKAGE DISEQUILIBRIUM

Linkage disequilibrium (LD) studies were performed using DNA from a population sample of neuropsychiatric disorder (BP-I) patients. The population sample and LD techniques were as described in Freimer et al., 1996, Nature Genetics 12:436-441. The present LD study took advantage of the additional physical markers identified via the physical mapping techniques described below.

6.1.2. YEAST ARTIFICIAL CHROMOSOME (YAC) MAPPING

For physical mapping, yeast artificial chromosomes (YACs) containing human sequences were mapped to the region being analyzed based on publicly available maps (Cohen *et al.*, 1993, C.R. Acad. Sci. 316:1484-1488). The YACs were then ordered and contig reconstructed by performing standard short tag sequence (STS)-content mapping with microsatellite markers and non-polymorphic STSs available from databases that surround the genetically defined candidate region.

6.1.3. BACTERIAL ARTIFICIAL CHROMOSOME (BAC) MAPPING

STSs from the short arm of human chromosome 18 were used to screen a human BAC library (Research Genetics, Huntsville, AL). The ends of the BACs were

cloned or directly sequenced. The end sequences were used to amplify the next overlapping BACs. From each BAC, additional microsatellites were identified. Specifically, random sheared libraries were prepared from overlapping BACs within the defined genetic interval. BAC DNA was sheared with a nebulizer (CIS-US Inc., Bedford, MA). Fragments in the size range of 600 to 1,000 bp were utilized for the sublibrary production. Microsatellite sequences from the sublibraries were identified by corresponding microsatellite probes. Sequences around such repeats were obtained to enable development of PCR primers for genomic DNA.

6.1.4. RADIATION HYBRID (RH) MAPPING

Standard RH mapping techniques were applied to a Stanford G3 RH mapping panel (Research Genetics, Huntsville, AL) to order all microsatellite markers and non-polymorphic STSs in the region being analyzed.

6.1.5. SAMPLE SEQUENCING: BASE-PAIR PERFECT SEQUENCING

Random sheared libraries were made from all the BACs within the defined genetic region. Approximately 9,000 subclones within the approximately 340 kb region containing the BAD interval were sequenced with vector primers in order to achieve an 8-fold sequence coverage of the region. All sequences were processed through an automated sequence analysis pipeline that assessed quality, removed vector sequences and masked repetitive sequences. The resulting sequences were then compared to public DNA and protein databases using BLAST algorithms (Altschul, *et al.*, 1990, J. Molec. Biol., 215:403-410).

All sequences were contiged using Sequencher 3.0 (Gene Code Corp.) and PHRED and PHRAP (Phill Green, Washington University) into a single DNA fragment of 340 kb.

6.2. RESULTS

Genetic regions involved in bipolar affective disorder (BAD) human genes had previously been reported to map to portions of the long (18q) and short (18p) arms of human chromosome 18 (Freimer *et al.*, 1996, Neuropsychiat. Genet. 67:254-263; and Freimer *et al.*, 1996, Nature Genetics 12:436-441).

High resolution physical mapping using YAC, BAC and RH techniques. In order to provide the precise order of genetic markers necessary for linkage and LD mapping, and to guide new microsatellite marker development for finer mapping, a high resolution physical map of the 18p candidate region was developed using YAC, BAC and RH techniques.

For such physical mapping, first, YACs were mapped to the chromosome 18 region being analyzed. Using the mapped YAC contig as a framework, the region from publicly available markers spanning the 18p region were also mapped and contiged with BACs. Sublibraries from the contiged BACs were constructed, from which microsatellite marker sequences were identified and sequenced.

To ensure development of an accurate physical map, the radiation hybrid (RH) mapping technique was independently applied to the region being analyzed. RH was used to order all microsatellite markers and non-polymorphic STSs in the region. Thus, the high resolution physical map ultimately constructed was obtained using data from RH mapping and STS-content mapping.

Linkage Disequilibrium. Prior to attempting to identify gene sequences, studies were performed to further narrow the neuropsychiatric disorder region. Specifically, a linkage disequilibrium (LD) analysis was performed using population samples and techniques as described in Section 6.1, above, which took advantage of the additional physical markers identified via the physical mapping techniques described below.

Initial LD analysis narrowed the interval which associates with BAD disorders to a 310 kb region of 18p. BAC clones within this newly identified neuropsychiatric disorder region were analyzed to identify specific genes within the region. A combination of sample sequencing, cDNA selection and transcription mapping analyses were combined to arrange sequences into tentative transcription units, that is, tentatively delineating the coding sequences of genes within this genomic region of interest.

Subsequent LD analyses further narrowed the BAD region of 18p to a narrow interval of approximately 25 kb. This was accomplished by identifying the maximum haplotype shared among affected individuals.

This newly identified narrow interval was found to map completely within one of the transcription units identified as described above. The gene corresponding to this transcription unit is referred to herein as the *HKNG1* gene. Thus, the results of the mapping analyses presented in this Section demonstrate that the *HKNG1* gene of human chromosome 18 is associated the neuropsychiatric disorder BAD.

Analysis of the BAD interval indicated that the 25 kb BAD disease-associated chromosomal interval identified in the linkage disequilibrium studies is contained within a 60 kb genomic region which contains a sequence referred to as GS4642 or rod photoreceptor protein (RPP) gene (Shimizu-Matsumoto, A. et al., 1997, Invest. Ophthalmol. Vis. Sci. 38:2576-2585).

7. EXAMPLE: SEQUENCE AND CHARACTERIZATION OF THE *HKNG1* GENE

As demonstrated in the Example presented in Section 6, above, the *HKNG1* gene is involved in the neuropsychiatric disorder BAD. The results presented in this Section further characterize the *HKNG1* gene and gene product. In particular, isolation of additional cDNA clones and analyses of genomic and cDNA sequences have revealed both the full length *HKNG1* amino acid sequence and the *HKNG1* genomic intron/exon structure. In particular, the nucleotide and predicted amino acid sequence of the *HKNG1* gene identified by these analyses disclose new *HKNG1* exon sequences, including new *HKNG1* protein coding sequence, discovered herein. Further, the expression of *HKNG1* in human tissue, especially neural tissue, is characterized by Northern and in situ hybridization analysis. The results presented herein are consistent with the *HKNG1* gene being a gene which mediates neuropsychiatric disorders such as BAD.

7.1. MATERIALS AND METHODS

HKNG1 cDNA Clone Isolation: Radiation hybridization of a human brain cDNA library was performed according to standard techniques and identified a full-length

HKNG1 cDNA clone. In addition, a *HKNG1* cDNA derived from a splice variant was isolated, as described in Section 7.2, below.

Northern Blot Analysis: Standard RNA isolation techniques and Northern blotting procedures were followed. The *HKNG1* probe utilized corresponds to the complementary sequence of base pairs 1367 to 1578 of the full length *HKNG1* cDNA sequence (SEQ ID NO. 1). Clonetech multiple tissue northern blots were probed. In particular, Clonetech human I, human II, human III, human fetal II, human brain II and human brain III blots were utilized for this study.

In Situ Hybridization Analysis: Standard in situ hybridization techniques were utilized. The *HKNG1* probe utilized corresponds to the complementary sequence of base pairs 910 to 1422 of the full length *HKNG1* cDNA sequence (SEQ ID NO. 1). Brains for in situ hybridization analysis were obtained from McLean Hospital (The Harvard Brain Tissue Resource Center, Belmont, MA 02178).

Other techniques: The remaining techniques described in Section 7.2, below, were performed according to standard techniques or as discussed in Section 6.1, above.

7.2. RESULTS

7.2.1. *HKNG1* Nucleotide and Amino Acid Sequence

A human brain cDNA library was screened and a full-length clone of *HKNG1* was isolated from this library, as described above. By comparing the isolated cDNA sequence to sequences in the public databases, a clone was identified which had been previously identified as GS4642, or rod photoreceptor protein (RPR) gene (GenBank Accession No. D63813; Shimizu-Matsumoto, A. et al., 1997, Invest. Ophthalmol. Vis. Sci. 38:2576-2585). Although Shimizu-Matsumoto et al. refer to GS4642 as a full-length cDNA sequence, the isolated *HKNG1* cDNA extends approximately 200 bp beyond the 5'end of the identified G34642 clone.

Importantly, the *HKNG1* clone isolated herein reveals that, contrary to the amino acid sequence described in Shimizu-Matsumoto et al., the full length *HKNG1* amino acid sequence contains an additional 29 amino acid residues N-terminal to what had previously been identified as the full-length *RPR* polypeptide. The full-length *HKNG1* nucleotide sequence (SEQ ID NO: 1) and the derived amino acid sequence of the full-length

HKNG1 polypeptide (SEQ ID NO: 2) encoded by this sequence are depicted in FIG. 1A-1B.

Full-length *HKNG1* cDNA sequence was compared with the genomic contig completed by random sheared library sequencing. Exon-intron boundaries were identified manually by aligning the two sequences in Sequencher 3.0 and by observing the conservative splicing sites where the alignments ended. This sequence comparison revealed that the additional cDNA sequence discovered through isolation of the full-length *HKNG1* cDNA clone actually belongs within three *HKNG1* exons.

Prior to the isolation and analysis of *HKNG1* cDNA described herein, nine exons were predicted to be present within the corresponding genomic sequence. As discovered herein, however, the *HKNG1* gene, in contrast, actually contains 11 exons, with the new cDNA containing sequence which corresponds to a new exon 1, exon 2 and a 5' extension of what had previously been designated exon 1. The genomic sequence and intron/exon structure of the *HKNG1* gene is shown in FIG. 3A-3R.

The breakdown of exons was confirmed by the perfect alignment of the cDNA sequence with the genomic sequence and by observation of expected splicing sites flanking each of the additional, newly discovered exons.

HKNG1 nucleotide sequence was used to search public EST databases. This search identified a putative *HKNG1* EST which was utilized to obtain a full length cDNA clone from the IMAGE Consortium library. Sequence of the isolated cDNA revealed that this new clone was derived from an *HKNG1* alternatively spliced *HKNG1* mRNA variant. In particular, this *HKNG1* variant is deleted for exon 3 of the full length 11 exon *HKNG1* sequence. The nucleotide sequence of this *HKNG1* variant (SEQ ID NO: 3) is depicted in FIG. 2A-2B. The amino acid sequence encoded by the *HKNG1* variant (SEQ ID NO: 4) is also shown in FIG. 2A-2B.

7.2.2. *HKNG1* GENE EXPRESSION

HKNG1 gene expression was examined by Northern blot analysis in various human tissues. A transcript of approximately 2 kb was detected in fetal brain, lung and kidney, and in adult brain, kidney, pancreas, prostate, testis, ovary, stomach, thyroid, spinal cord, lymph node and trachea. An approximately 1.5 kb transcript was also seen in trachea.

In addition, a larger transcript of approximately 5 kb was detected in all adult neural regions tested (that is, cerebellum, cortex, medulla, spinal cord, occipital pole, frontal lobe, temporal, putamen, amygdala, caudatte nucleus, corpus callosum, hippocampus, whole brain, substantia nigra, subthalamic nucleus and thalamus). Once again, this is in direct contrast to previous Northern analysis of the *RPR* gene, which reported that expression was limited to the retina (Shimizu-Matsumoto, A. et al., 1997, Invest. Ophthalmal. Vis. Sci. 38:2576-2585).

Analysis of *HKNG1* the tissue distribution was extended through an in situ hybridization analysis. In particular, the *HKNG1* mRNA distribution in normal human brain tissue was analyzed. The results of this analysis are depicted in FIG. 4. As summarized in FIG. 4, *HKNG1* is expressed throughout the brain, with transcripts being localized to neuronal and grey matter cell types.

Finally, expression of *HKNG1* in recombinant cells demonstrates that the *HKNG1* gene encodes a secreted polypeptide(s).

8. A MISSENSE MUTATION WITHIN *HKNG1* CORRELATES WITH BAD

The Example presented in Section 6, above, shows that the BAD disorder maps to an interval completely contained within the *HKNG1* gene of the short arm of human chromosome 18. The Example presented in Section 7, above, characterizes the *HKNG1* gene and gene products. The results presented in this Example further these studies by identifying a mutation within the coding region of a *HKNG1* allele of an individual exhibiting a BAD disorder.

Thus, the results described herein demonstrate a positive correlation between a mutation which encodes a non-wild type *HKNG1* polypeptide and the appearance of the neuropsychiatric disorder BAD. The results presented herein, coupled with the results presented in Section 6, above, identify *HKNG1* as a gene which mediates neuropsychiatric disorders such as BAD.

8.1. MATERIALS AND METHODS

Pairs of PCR primers that flank each exon (see TABLE 1, above) were made and used to PCR amplify genomic DNA isolated from BAD affected and normal individuals. The amplified PCR products were analyzed using SSCP gel electrophoresis or by DNA sequencing. The DNA sequences and SSCP patterns of the affected and controls were compared and variations were further analyzed.

8.2. RESULTS

In order to more definitively show that the *HKNG1* gene mediates neuropsychiatric disorders, in particular BAD, a study was conducted to explore whether a *HKNG1* mutation that correlates with BAD could be identified.

First, exon scanning was performed on all eleven exons of the *HKNG1* gene using chromosomes isolated from three affected and one normal individual from the Costa Rican population utilized for the LD studies discussed in Section 6, above. No obvious mutations correlating with BAD were found through this analysis.

Next, *HKNG1* intron and 3'-untranslated regions within the 25 kb BAD interval were scanned by sscp and/or sequencing for all variants among three affected and one normal individual from the same population. More than 60 variants were identified after scanning approximately two-thirds of the 25 kb genomic interval, which can be genotyped and analyzed by haplotype sharing and LD analyses, as described above, in order to identify ones which correlate with bipolar affective disorder. Fig. 5 lists selected variants identified through this study.

Exon scanning using chromosomal DNA from the general population, however, successfully identified a *HKNG1* missense mutation in an individual affected with BAD who did not share the common diseased haplotype identified by the LD analysis provided above. In particular, exon scanning was done on exons 1-11 of *HKNG1* nucleic acid from 129 individuals from the general population affected with BAD.

This analysis identified a point mutation in the coding region of exon 7 not seen in non-bipolar affected disorder individuals. Specifically, the guanine corresponding to nucleotide residue 604 of SEQ ID NO:1 (or nucleotide residue 550 of SEQ ID NO:3) had mutated to an adenine. *HKNG1* protein expressed from this mutated *HKNG1* allele

comprises the substitution of a lysine residue at amino acid residue 202 of SEQ ID NO:2 (or amino acid residue 184 of SEQ ID NO:4) in place of the wild type glutamic acid residue.

Additional *HKNG1* polymorphisms relative to the consensus *HKNG1* wild type sequence, and which, therefore, represent *HKNG1* alleles, were identified through sequence analysis of the *HKNG1* alleles within the NIMH schizophrenia collection. These variants are depicted in FIG. 5.

9. IDENTIFICATION OF A SECOND CAUSATIVE GENE FOR BIPOLAR AFFECTIVE DISORDER

The data presented in Section 9 above implicate LRP-2 and RAP in the etiology of bipolar affective disorder and schizophrenia. Gene mapping data on RAP was therefore examined. RAP maps to the short arm of human chromosome 4 at position 4p16.3 (Korenberg *et al.*, 1994, *Genomics 22*:88-93). To further refine the location of the RAP gene, a Radiation Hybrid database was searched on the World Wide Web. RAP was found to be located on human chromosome 4, about 39 cR3000 (or about 8 Mb) from the marker D4S394 (towards the telomere). A previous study has reported a very strong linkage of bipolar affective disorder to human chromosome 4 in the region of this marker with a 2-point lod score of 4.1 (Blackwood *et al.*, 1996, *Nature Genetics 12*:427-430). The linkage region includes the location of RAP. This information, combined with the data presented in Section 9 above, identify RAP as a causative gene in this locus responsible for bipolar affective disorder.

ins C10

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention, and functionally equivalent methods and components are within the scope of the invention. Indeed, various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and accompanying drawings.

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or

patent application was specifically and individually indicated to be incorporated by reference.